STOCHASTIC MODELING OF ECONOMIC VARIABLES FOR PENSION PLAN PROJECTIONS

by

Henry Yuen

B.Sc. (Hons.), Simon Fraser University, 2002

a Project submitted in partial fulfillment
of the requirements for the degree of
Master of Science
in the
Department of Statistics and Actuarial Science
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2011

APPROVAL

Name: Henry Yuen

Degree: Master of Science

Title of Project: Stochastic Modeling of Economic Variables for Pension Plan Projections

Examining Committee: Dr. Tim Swartz
Chair

_____

Dr. Gary Parker, Senior Supervisor

_____

Dr. Yi Lu, Supervisor

_____

Dr. Richard Lockhart, External Examiner

Date Approved: _____

Abstract

Key economic variables for pension plan projections are identified. These variables are modeled based on data since 1955. Several time series models are considered including regime switching models and a simplified Wilkie Model. To investigate the dynamics of different models, simulations are carried out to project the economic series for the next 50 years using different starting values. All the models for each series are then compared using different criteria including economic theories and common actuarial practice. The best model out of the considered models for each series is selected for pension plan projections. Using actual starting values, simulation is performed again to project the economic series to model a sample defined benefit (DB) pension plan and a sample defined contribution (DC) pension plan. The total employer contributions as a percentage of wages for the DB plan, and the replacement ratio for the DC plan are studied.

Keywords: Time Series; Parameter Estimation; Simulation; Pension Plan Projection; Employer Contributions; Replacement Ratio

To my family.

"In the darkest times, hope is something you give yourself.
That is the meaning of inner strength."

— *Avartar: The Last Airbender,* NICKELODEON, *2007*

Acknowledgments

I would like to thank my supervisor, Dr. Gary Parker, for his guidance over the last year and a half. He has been very generous and patient with me and other students. I am fortunate that I had the chance to work with him and learn from him, both as a researcher and an educator.

I would also like to thank the members of the examining committee, Dr. Yi Lu and Dr. Richard Lockhart, for giving me practical comments and suggestions for this project as well as future research.

I would like to express my thanks to the Department of Statistics and Actuarial Science for their support, in particular the professors whose lectures I attended and whose offices I frequently occupied. The professors are Dr. Tom Loughin, Dr. Steve Thompson, Dr. Cary Chi-Liang Tsai, and the whole examining committee. Also, I would like to thank Barbara Sanders for supporting me going back to school eight years after my undergraduate degree.

Moreover, I wish to thank the fellow graduate students for giving me a memorable journey and friendship during the pursuit of the degree, in particular Oksana Chrebtii and Andrew Henrey (his ninja themed presentations are especially memorable).

Last but not least, I am thankful that Eve Belmonte and Julia Viinikka agreed to write my reference letters for admission to the program. Also, I am grateful to my parents for their love and support as well as my partner for encouraging me to follow my interests, and, more importantly, putting up with all the 'side effects' of being stressed.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Pension Plan Projections

For a defined contribution (DC) pension plan, the employer contribution as a percentage of wages is usually fixed. The balance of the pension account at retirement is affected by the wage increases throughout the career of the employee. It is also heavily affected by the rates of return earned every year during employment. At retirement, one of the options the employee has is to purchase an annuity from an insurance company. The amount of pension converted from the lump sum is based on the prevailing interest rate at retirement. The replacement ratio, which is the ratio of annual pension amount to annual wages just before retirement, is influenced by all the economic variables mentioned above.

For a defined benefit (DB) pension plan, the amount of pension payable to employees at retirement is based on a predetermined formula. Also, auxiliary benefits (bridging benefits, early retirement subsidy, survivor benefits, and post-retirement indexing) specified in the plan text provide additional values to the pension. The employer is responsible for fully funding the pension plan, and therefore, is required to increase the contribution amount when the pension plan is under-funded. In case of a pension plan surplus, employer is usually allowed to take a contribution holiday (i.e. contribution is not required). Wage increases, interest rate for the valuation of the pension plan, and asset returns determine the funded status of a DB pension plan. Accounting rules in Canada require the employer to incorporate the pension expense, pension assets, and pension liabilities into the financial statements. Prevailing interest rates and the outlook of the future assets return at the time of financial reporting have a significant impact on the pension expense and pension liabilities.

From time to time, employer or employees may want to change the benefit provisions of the pension plan. Employer usually would like to lower the cost of maintaining the pension plan while the employees usually would like to increase the value of the pension payable at retirement. These changes in value, and thus costs and employer contributions, are heavily influenced by the economic variables.

In order to understand the relationships between the economic variables and the cost of a pension plan, a projection can be performed. A DB pension plan projection simulates the amount of pension assets and liabilities in the future. At each valuation date, employer contributions are revised to reflect the funding status of the pension plan. A DC pension plan projection simulates the balance of the pension account at retirement, and converts the balance into an annuity based on the simulated interest rate at retirement.

Traditionally, a pension plan projection uses deterministic variables which have the same values as in the last valuation. This method only provides a single point estimate of the changes. However, using stochastic modeling and simulation, a distribution of the results can be obtained instead. Consequently, the mean and standard deviation can be studied. Percentiles can also be calculated and are of great importance because they represent the 'good' and 'bad' scenarios. Another advantage of stochastic modeling is that correlated series can be modeled together to reflect the correlation between them. Actuaries have already employed such correlations in determining the values of the valuation variables; therefore, actuaries are familiar with those relationships, and expect them to be part of the modeling.

1.2 Economic Variables for Pension Plan Projections

DB pension plan modeling has been investigated in discrete time by Cairns and Parker (1997), Dufresne (1989), Haberman (1994) as well as in continuous time by Cairns (1996) with the rates of return modeled as a single stochastic variable. Zhang and Hou (2011) calculated the optimal investment strategy for a DC pension plan where the wage increase and inflation are stochastic. Chang and Cheng (2002) used stochastic interest rates and inflation rates to model a DB pension plan. A stochastic wage index was also calculated as a function of the inflation rates.

To perform a DB pension plan valuation in Canada, actuaries are required to determine the values of the economic and demographic variables. For economic variables, the future

wage increases, and the future return on the assets are needed. Usually, a building block approach is used to determine the wage increases. That is, future wage increase is determined as the sum of three components: future expected inflation, future general productivity increases, and future merit increases. The expected return on the assets is a weighted average of the returns of each asset class in the target asset mix. Most of the DB pension assets in Canada include fixed-income, Canadian equity, and global equity in the target asset mix. For DC pension plans, it is assumed that employee's pension account includes the same classes of asset as in the DB pension assets. Since a pension plan projection is a series of valuations at different points of time in the future, the economic variables needed for pension plan projection are the same as those needed for a pension plan valuation. Therefore, this report employs more economic variables than the ones mentioned in the articles above.

In this report, five different economic variables are identified based on the Canadian DB pension plan valuation requirements. They are:

- inflation;
- wage;
- long-term interest rate;
- Canadian equity return; and
- global equity return.

Note that all the demographic variables are assumed to be deterministic in this report.

1.3 Time Series Models

Box and Jenkins (1976) proposed using the autoregressive integrated moving average (ARIMA) model to model time series. Since then, a lot of variations of the model have been created. Box and Jenkins' model assumed constant innovation variance while Engle (1982) proposed a model with non-constant variances conditional on the past. Engle's model is called autoregressive conditional heteroscedasticity (ARCH) model.

Hamilton (1989) considered a nonlinear stationary process. He proposed a model in which the trend of a series changes in response to discrete unobserved events. In other words, the time series follows an ARIMA model but the trend parameter changes when

the underlying regime changes. The transition of the regime from one state to another is assumed to follow a Markov chain. This model is generally called a regime switching model.

Since the publication of Hamilton's paper, regime switching models have been generalized to model the effect of the underlying regime on different parameters in ARIMA and ARCH models (for example Hamilton and Susmel (1994)). In the mean time, Wilkie (1986, 1995) has proposed to model different economic series together through a cascade model. He first modeled inflation. Then he took advantage of the correlation between economic series to model them in layers, using one layer to explain another. For example, wages are modeled using inflation with the error terms following an autoregressive model of order one.

In this report, only the following models are considered:

- white noise model;
- autoregressive model of order one (AR(1));
- autoregressive conditional heteroscedasticity model of order one (ARCH(1));
- regime switching model with two regimes; and
- transfer function (Wilkie model).

For the ARCH(1) model, the series is modeled as an AR(1) model with non-constant variances based on the previous observation. Also, three versions of the ARCH(1) model are considered with different parameters in the variance components set to zero. For the regime switching model, the class of models is also restricted to the AR(1) and ARCH(1) models only. For the transfer function model, both one lag and two lags are considered.

Only the five models mentioned above are considered because these models are simple enough to implement the pension plan projections, and flexible enough to do an adequate job to model the dynamics of the economic variables. Simple models can help actuaries, employers, and employees understand the dynamics of economic series more easily. This is an important aspect because stochastic modeling is not common outside the academic sector, and therefore, simplicity will encourage the usage of time series models.

1.4 Data

With the exception of global equity return, all series identified above can be found in the CANSIM series of Statistics Canada. Global equity returns were obtained from the Morgan

Stanley Capital International (MSCI) website.

The data obtained is then studied, and, for some of the series, only the more recent data is used to estimate the parameters. The main reason for using only recent data is to better reflect the dynamics of the future by eliminating trends and volatility that only existed in the past.

1.5 Model Selection

In general, for an AR(1) model, the autocorrelation function (ACF) decreases slowly as the lag increases, and the partial autocorrelation function (PACF) drops to almost zero after lag one. If the ACF and PACF display the mentioned patterns, an AR(1) model is considered an appropriate model. If the ACF is only significant at lag zero and PACF is not significant at any lag, a white noise model is considered appropriate. See Appendices B.1 and B.2 for details of the white noise model and the AR(1) model respectively. The graphs of ACF and PACF are plotted for each series to determine the appropriate model.

Once the white noise or AR(1) model is fitted to the data, the residual process is studied. In particular, the cross-correlation function between the squared residuals and the series itself is plotted. If the cross-correlation function is significant at lag minus one, a larger residual in one period is likely to be followed by a larger observation in the next period. This suggests that the variance of the innovations might change as the series changes; therefore, using an ARCH(1) model might be appropriate. As mentioned above, three versions of the ARCH(1) model are fitted to the data. They are the full model (Appendix B.3), the non-centered model (Appendix B.4), and the proportional model (Appendix B.5).

For inflation and both equity returns, the next step is to consider regime switching models. If a certain model is not considered appropriate in the previous step, the regime switching version of the same model is also not considered in this step. Regime switching models are generally considered appropriate for economic series as Koop, Milas, and Osborn (2008) pointed out that "nonlinear time series models in general, and regime-switching models in particular, have increased our understanding of many issues in economics and finance". See Appendices B.7-B.10 for details of the regime switching models. Note that we have assumed the regime for the equity returns is independent of the regime for inflation. To model the Canadian equity return and global equity return being influenced by the same regime, a vector white noise model is used. See Appendix B.6 for details of this model.

For wages and long-term interest rate, the transfer function model is considered. In particular, inflation is used as an exogenous variable to explain both series. The error terms are assumed to follow an AR(1) model. Both transfer function models for two lags (Appendix B.11) and one lag (Appendix B.12) are considered for comparison purposes. Note that for wages, this is the model proposed by Wilkie (1995). However, Wilkie proposed a more complex version for long-term interest rates which involves inflation, two autoregressive series, and the error term from the dividend yield. This version is not considered in this report partly because the dividend yield is not modeled.

After all the appropriate models are selected, parameters are estimated using the maximum likelihood method. The asymptotic standard deviations of the parameters are also calculated when the Fisher information matrix is known. The parameter estimates are then used to project the series for the next 50 years using simulation (100,000 trials each). For comparison purposes, different starting values are used: one close to the long-term mean, and one far from the long-term mean. Mean, standard deviation, and percentiles are calculated based on the projected values. Other information from the simulation is also calculated if it is useful in identifying the final model.

After all the above steps are done, the models are compared for each series. The final model is selected based on the parameter estimates, the reasonableness of the dynamics implied by the model, extreme values in the projection, economic theories, and common actuarial practices.

The formulas of the log-likelihood function, first and second derivative of the log-likelihood, and the long-term mean and long-term variance for selected models are summarized in Appendix B.

1.6 Outline

Chapters 2, 3, 4, 5 provide the details of the model selection for inflation, wages, long-term interest rate, and Canadian equity returns and global equity returns respectively. Chapter 6 provides the methodology of the pension plan projection and two simple illustrations. Chapter 7 provides a brief conclusion.

Chapter 2

Inflation

2.1 Data

Annual consumer price index (CPI) is available from 1914 to 2010 based on CANSIM series V41693271. The consumer price index is calculated using the 2005 basket of all items (i.e. not excluding any items) across Canada.

The force of annual inflation is calculated as follow:

$$X_I(t) = ln\frac{CPI(t)}{CPI(t-1)}. \tag{2.1.1}$$

In Figure 2.1, the force of annual inflation is significantly more volatile on the left side of the vertical line (year 1955) than on the right side. Note that Bank of Canada adopted an inflation-control target in 1991, and revised the target in November 2011. The target "aims to keep total CPI inflation at the 2 per cent midpoint of a target range of 1 to 3 per cent over the medium term" ("Inflation-Control Target", n.d.). As a result, it is reasonable to believe that big fluctuations in inflation are not expected in the future in Canada, and consequently, only data from 1955 onwards is used to estimate the parameters.

2.2 Model Selection

2.2.1 AR(1) Model

The ACF and PACF in Figure 2.2 show the pattern of an AR(1) series. As a result, an AR(1) model is appropriate for the force of annual inflation.

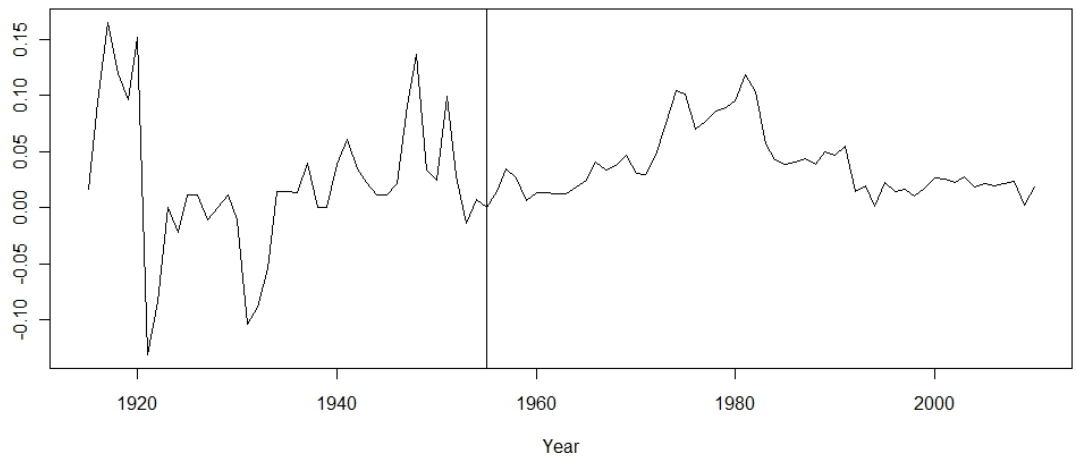

Figure 2.1: Force of Annual Inflation from 1915 to 2010

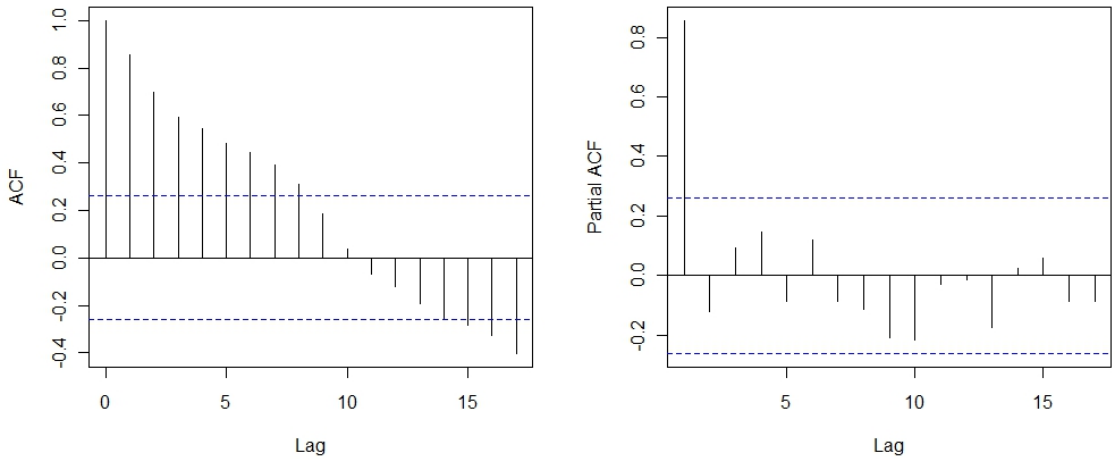

Figure 2.2: Autocorrelation Function and Partial Autocorrelation Function for the Force of Annual Inflation

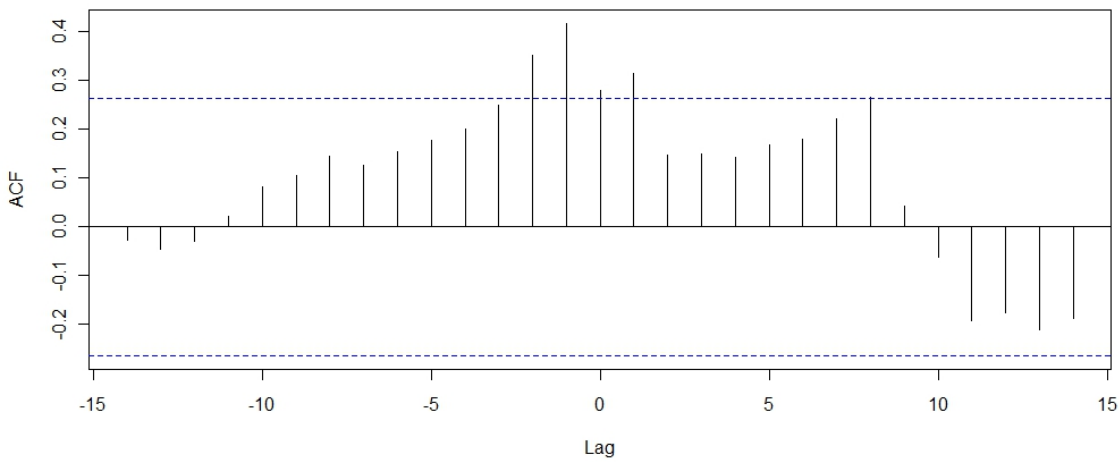

Figure 2.3: Cross-correlation Function for the Force of Annual Inflation and its Squared AR(1) Residual

2.2.2 ARCH(1) Models

The parameter estimates for the AR(1) model are summarized in Table 2.3. The residual is then calculated based on the parameter estimates. Since there is significant cross-correlation at lag minus one between the series and the squared residual, an ARCH(1) model is considered. Note that the cross-correlations at lag minus two, zero, and one are also significant. Although a significant cross-correlation at lag minus two suggests an ARCH(2) model, this is not within the class of models considered. Significant cross-correlation at lag zero and one imply the observation at current and next time period, respectively, affect the volatility at current time period. This is not useful for the model because current and future observations cannot be used to formulate the current volatility. Only past observations are available to calculate the volatility in the current period.

Note that only the full ARCH(1) model (ARCH(F)) and the non-centered ARCH(1) model (ARCH(N)) can be used. Since the first observation is zero, the variance for the second observation is zero under the proportional ARCH(1) model (ARCH(P)). Consequently, the log-likelihood is undefined. However, as stated in Section 2.2.3, the parameter estimates of the regime switching ARCH(N) model suggest a regime switching ARCH(P) model. Therefore, the ARCH(P) model is also considered with the first observation removed.

2.2.3 Regime Switching AR(1) and ARCH(1) Models

For the force of annual inflation, regime switching models are also considered, namely: a regime switching AR(1) model (RSAR(1)), a regime switching ARCH(F) model (RSARCH(F)), a regime switching ARCH(N) model (RSARCH(N)), and a regime switching ARCH(P) model (RSARCH(P)).

The parameter estimates for the RSARCH(F) model (the first column of estimates in Table 2.1) do not produce a stationary standard deviation. Specifically, regime two is not a stationary process since $\hat{\alpha}_{I2}^2 + \hat{\gamma}_{I2}^2 > 1$. Consequently, this model does not produce meaningful results for pension plan projection. However, a search in the neighbourhood of the maximum likelihood estimates is performed to find a parameter set that produces a stationary standard deviation. This new parameter set, inevitably, does not maximize the log-likelihood function but should be chosen only if the new log-likelihood value is reasonably close to the maximum value. The standard deviation for regime two is,

$$SD_{I2} = \frac{\hat{\lambda}_{I2}}{\sqrt{1 - \hat{\alpha}_{I2}^2 - \hat{\gamma}_{I2}^2}} = \frac{0.0096}{\sqrt{1 - 0.9959^2 - 0.3358^2}} = \frac{0.0096}{\sqrt{-0.1046}}.$$

Since changing the value of $\hat{\gamma}_{I2}$ decreases the log-likelihood value significantly, the value of $\hat{\gamma}_{I2}$ remains unchanged at 0.3358. As a result, $\hat{\alpha}_{I2}$ needs to be less than 0.9419 in order to have a stationary process, and therefore, a step of 0.06 ($\approx 0.9959 - 0.9419$) is used. Table 2.1 summarizes different log-likelihood values (LLV) for different parameter sets. The last column of the table is the parameter set with the highest log-likelihood value among all the parameter sets that produce a stationary standard deviation. Note that negative values of $\hat{\gamma}_{I1}$ have the same effect as positive values.

For RSARCH(N), both $\hat{\lambda}_{I1}$ and $\hat{\lambda}_{I2}$ attained the minimum boundary of the maximization, which is 0.0001. This suggests that an RSARCH(P) might be appropriate. However, as mentioned in Section 2.2.2, the first observation has to be removed so that the log-likelihood function is defined. Note that the maximum likelihood estimates for RSARCH(N) and RSARCH(P) are different and not comparable because the data used for the maximization is different.

The estimate of α_{I1} in RSARCH(P) is essentially zero (1e-8). This implies one regime is a white noise process while the other regime is an AR(1) process. In order for the model to be an ARCH(P) model in both regimes, a search is also performed for this set of parameters.

Table 2.1: Search for Stationary Standard Deviation for RSARCH(F) Model for the Force of Annual Inflation

$\hat{\mu}_{I1}$	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177
$\hat{\alpha}_{I1}$	0.0360	0.0360	0.0360	**0.0960**	**0.0960**	0.0360	0.0360
$\hat{\lambda}_{I1}$	0.0074	0.0074	0.0074	0.0074	0.0074	0.0074	0.0074
$\hat{\gamma}_{I1}$	0.0011	**-0.0589**	**0.0611**	0.0011	**0.0611**	0.0011	0.0011
$\hat{\mu}_{I2}$	0.0304	0.0304	0.0304	0.0304	0.0304	0.0304	0.0304
$\hat{\alpha}_{I2}$	0.9959	0.9959	0.9959	0.9959	0.9959	**1.0559**	**0.9359**
$\hat{\lambda}_{I2}$	0.0096	0.0096	0.0096	0.0096	0.0096	0.0096	0.0096
$\hat{\gamma}_{I2}$	0.3358	0.3358	0.3358	0.3358	0.3358	0.3358	0.3358
$\hat{\rho}_{I12}$	0.0412	0.0412	0.0412	0.0412	0.0412	0.0412	0.0412
$\hat{\rho}_{I21}$	0.0423	0.0423	0.0423	0.0423	0.0423	0.0423	0.0423
LLV	167.5427	167.5090	167.5065	167.4686	167.4360	167.3784	167.3768

A step of 0.1 is used. The final parameter set is chosen based on the highest log-likelihood value. Table 2.2 summarizes different log-likelihood values for different parameter sets. The first and last columns of estimates in the table denote the initial and final parameter sets.

Table 2.2: Search for true RSARCH(P) Model for the Force of Annual Inflation

$\hat{\mu}_{I1}$	0.0178	0.0178	0.0178	0.0178	0.0178	0.0178	0.0178
$\hat{\alpha}_{I1}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**0.1000**
$\hat{\gamma}_{I1}$	0.4995	0.4995	**0.5995**	**0.3995**	**0.3995**	**0.3995**	**0.5995**
$\hat{\mu}_{I2}$	0.0361	0.0361	0.0361	0.0361	0.0361	0.0361	0.0361
$\hat{\alpha}_{I2}$	0.8539	0.8539	0.8539	0.8539	0.8539	0.8539	0.8539
$\hat{\gamma}_{I2}$	0.2920	0.2920	0.2920	0.2920	0.2920	0.2920	0.2920
$\hat{\rho}_{I12}$	0.3786	**0.2786**	0.3786	0.3786	**0.2786**	**0.4786**	0.3786
$\hat{\rho}_{I21}$	0.0648	0.0648	0.0648	0.0648	0.0648	0.0648	0.0648
LLV	168.1118	168.0157	167.9723	167.8335	167.8075	167.6976	167.3863

2.3 Parameters Estimation

The final parameter estimates as well as the long-term mean (LTM) and long-term standard deviation (LTSD) are summarized in Tables 2.3 and 2.4. The asymptotic standard error of the estimation is shown in brackets, if applicable.

Table 2.3: AR(1)/ARCH(1) Models Parameter Estimates for the Force of Annual Inflation

Model	AR(1)	ARCH(F)	ARCH(N)	ARCH(P)
$\hat{\mu}_I$	0.0405 (0.0139)	0.0218 (0.0065)	0.0353 (0.0090)	0.0269 (0.0029)
$\hat{\alpha}_I$	0.8624 (0.0652)	0.8794 (0.0906)	0.7984 (0.0852)	0.2429 (0.0891)
$\hat{\sigma}_I$	0.0141 (0.0013)	N/A	N/A	N/A
$\hat{\lambda}_I$	N/A	0.0097 (0.0014)	0.0080 (0.0000)	N/A
$\hat{\gamma}_I$	N/A	0.3570 (0.0933)	0.2613 (0.0463)	0.5850 (0.0563)
LTM	0.0405	0.0218	0.0353	0.0269
LTSD	0.0278	0.0307	0.0225	0.0203

Table 2.4: Regime Switching AR(1)/ARCH(1) Models Parameter Estimates for the Force of Annual Inflation

Model	RSAR(1)	RSARCH(F)	RSARCH(N)	RSARCH(P)
$\hat{\mu}_{I1}$	0.0179	0.0177	0.0184	0.0178
$\hat{\alpha}_{I1}$	0.0817	0.0360	0.2338	0.1000
$\hat{\sigma}_{I1}$	0.0073	N/A	N/A	N/A
$\hat{\lambda}_{I1}$	N/A	0.0074	0.0001	N/A
$\hat{\gamma}_{I1}$	N/A	0.0011	0.9026	0.5995
$\hat{\mu}_{I2}$	0.0456	0.0304	0.0366	0.0361
$\hat{\alpha}_{I2}$	0.8560	0.9359	0.8492	0.8539
$\hat{\sigma}_{I2}$	0.0158	N/A	N/A	N/A
$\hat{\lambda}_{I2}$	N/A	0.0096	0.0001	N/A
$\hat{\gamma}_{I2}$	N/A	0.3358	0.2951	0.2920
$\hat{\rho}_{I12}$	0.0356	0.0412	0.4219	0.3786
$\hat{\rho}_{I21}$	0.0405	0.0423	0.0553	0.0648
LTM	0.0308	0.0240	0.0345	0.0335
LTSD	0.0248	0.0308	0.0312	0.0237

2.4 Time Series Projection

For the force of annual inflation, simulations are only performed for the regime switching models. Two initial values are used: 3% (close to the long-term mean) and 10% (far from the long-term mean). Figures 2.4, 2.5, 2.6, and 2.7 show the mean, standard deviation, 5th and 95th percentiles, and the minimum and maximum of the projection data respectively.

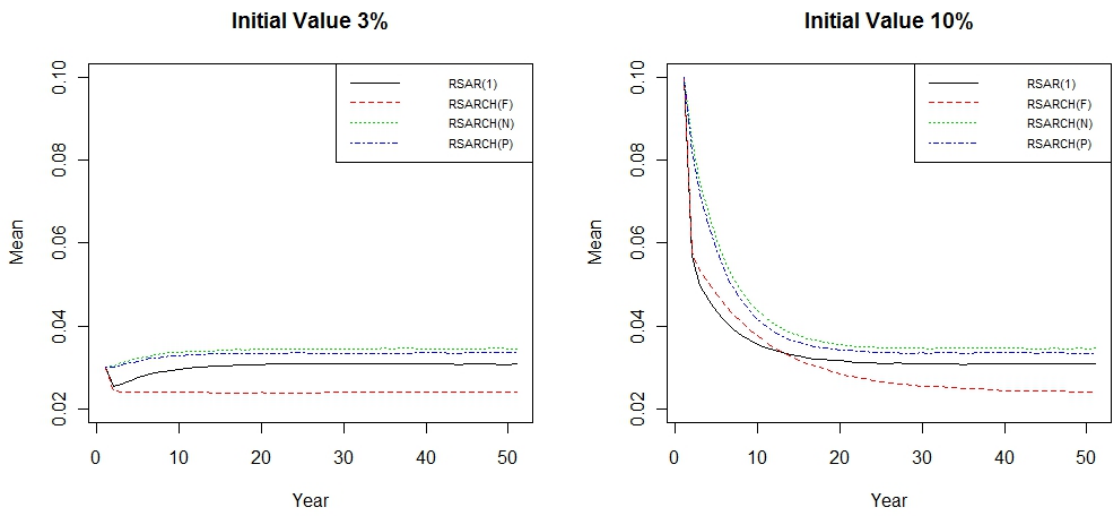

Figure 2.4: Mean of Projection for the Force of Annual Inflation

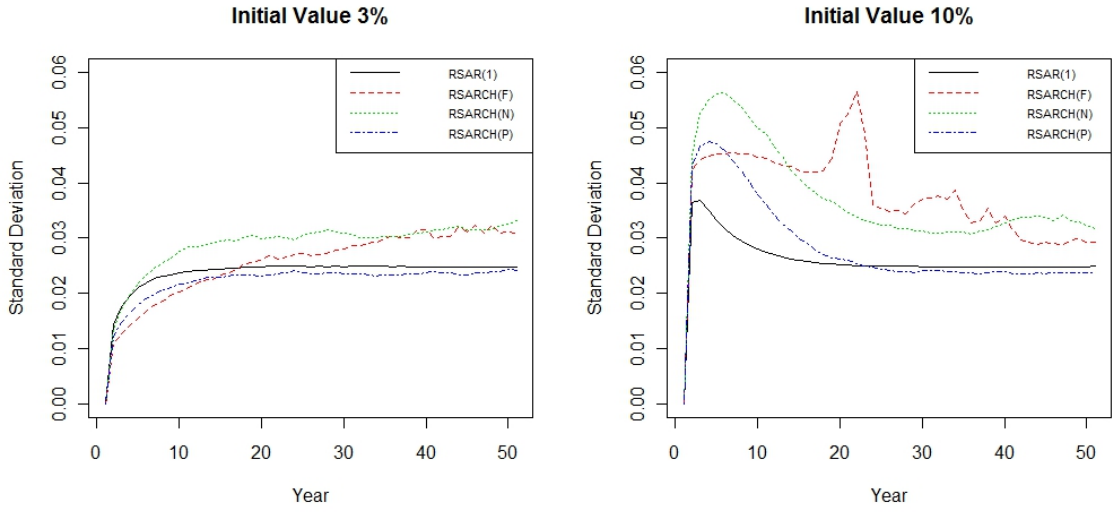

Figure 2.5: Standard Deviation of Projection for the Force of Annual Inflation

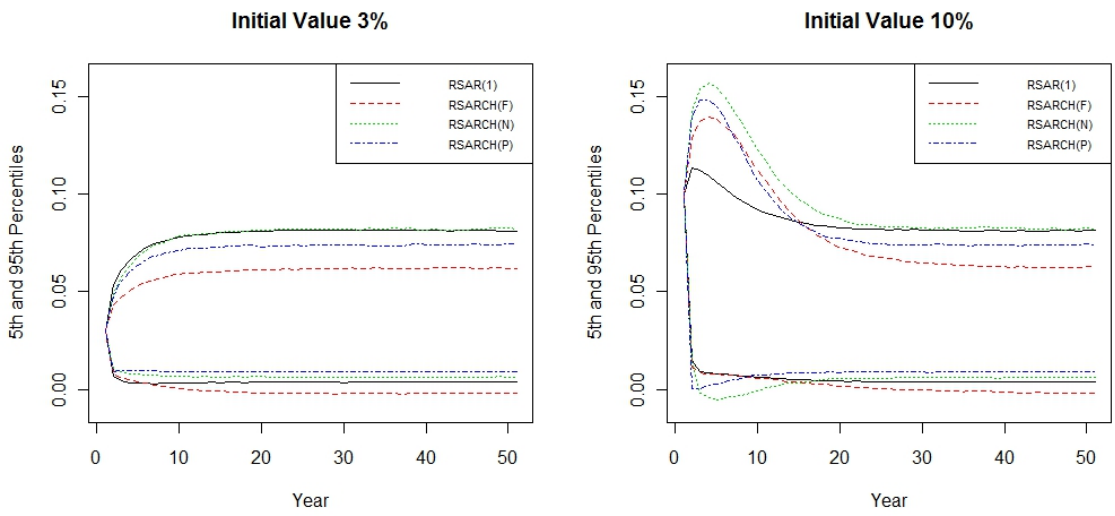

Figure 2.6: 5th and 95th Percentiles of Projection for the Force of Annual Inflation

Figure 2.7: Minimum and Maximum of Projection for the Force of Annual Inflation

2.5 Final Model

Comparing the parameter estimates for the AR(1) and ARCH(1) models, they are very different. For the ARCH(F) model, the long-term mean μ_I is also used to determine the variance of the innovations. As a result, $\hat{\mu}_I$ is significantly lower because it has to balance between the mean portion as well as the variance portion of the log-likelihood function to produce the maximum. Note that the ARCH(F) model can be re-parameterized by assuming a mean for the variance of the innovations that is different than the long-term mean; however, given the data is modeled using an AR(1) model, it is difficult to justify and interpret such model.

The ARCH(P) model also has a significantly lower mean. This is because the local volatility is very large (as indicated by the large value of $\hat{\gamma}_I$). Consequently, the larger observations are viewed as large fluctuations around a lower mean instead of small fluctuations around a higher mean. Moreover, the long-term standard deviation for the force of annual inflation is between 0.020 and 0.032 for all models. Therefore, in order to preserve the long-term volatility suggested by the data, the autocorrelation $\hat{\alpha}_I$ is also significantly lower to balance out the large value of $\hat{\gamma}_I$.

Although there is no graphs or tables to demonstrate the appropriateness of the regime switching models, regime switching models provide more flexibility to model inflation. Therefore, all the non-regime switching models are not considered here, unless the parameters estimation and/or dynamics for all the regime switching models are unsatisfactory.

The RSAR(1) model produced reasonable simulation data and dynamics; however, compared to the ARCH models, the RSAR(1) model does not have the feature of conditional heteroscedasticity, which has been illustrated in Figure 2.3.

For the regime switching models, the mean and the 5th and 95th percentiles look reasonable in Figures 2.4 and 2.6. However, Figure 2.5 shows that the standard deviations for both RSARCH(F) and RSARCH(N) models do not converge to a long-term value (i.e. the process is not stationary) even after 50 years. Although the theoretical long-term standard deviations exist for both models, the local volatility is still large enough to produce a very large number. This is illustrated in Figure 2.7. The range between the minimum and the maximum projected value is largest for the RSARCH(F) model, then followed by the RSARCH(N) model. The extreme values in these two models are so large that they move the standard deviation away from the stationary value. For pension plan projections,

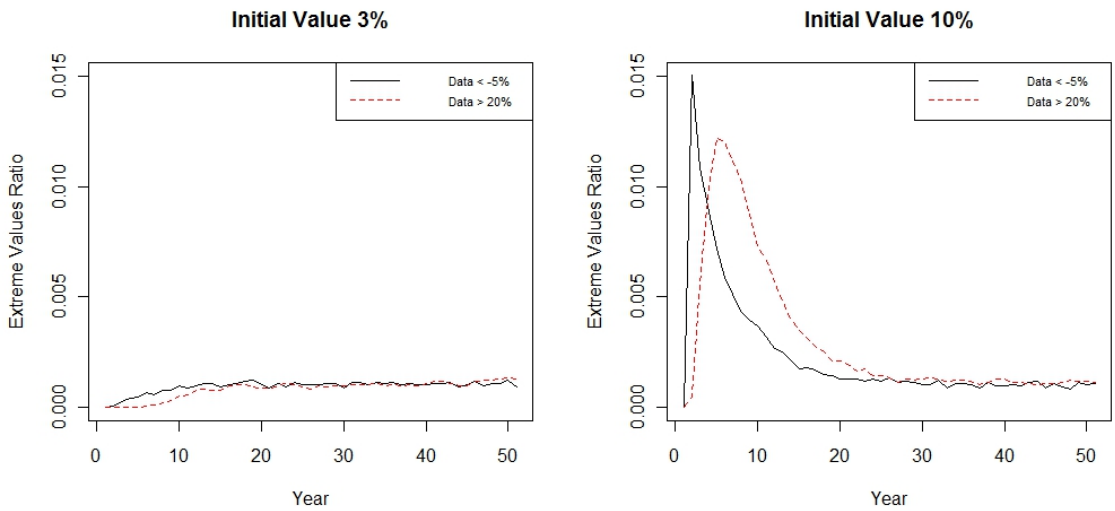

Figure 2.8: Extreme Value Ratios of Projection for the Force of Annual Inflation RSARCH(P) Model

extreme values do not produce meaningful conclusions; therefore, both RSARCH(F) and RSARCH(N) models should be used with caution when projecting pension plans.

For the RSARCH(P) model, the range of the simulated data is not large enough to produce a non-stationary series, which is an advantage over the RSARCH(F) and RSARCH(N) models. To ensure the usability of the RSARCH(P) model, an investigation of the extent of the extreme values is needed. The extreme values are defined as forces of annual inflation greater than 20% or smaller than -5%. The extreme value thresholds are not symmetric because, with these models, deflation is less likely to occur than high inflation. Figure 2.8 shows the extreme value ratio (number of extreme values divided by 100,000) for the RSARCH(P) model. Out of 100,000 trials, about 0.1% of the time the simulated data is over 20% or under -5% when the series becomes stationary. For an initial value of 10%, extreme values, as expected, are more common for the first 10 years. Although extreme values are not good for pension plan projections, the frequency of its occurrence shown in Figure 2.8 is acceptable.

In conclusion, RSARCH(P) model is the most appropriate model for the force of annual inflation for pension plan projections.

Chapter 3

Wage Index

3.1 Data

Average weekly earnings (AWE) are available at Statistics Canada from 1939 to 2010 based on the following series:

1. Historical Statistics of Canada (second edition) E49 series: This series consists of annual AWE across Canada from 1939 to 1975. For data between 1939 and 1961, the 1948 Standard Industrial Classification (SIC) was used while data between 1957 and 1975 used the 1960 SIC.

2. CANSIM series V75249: This series consists of monthly AWE across Canada from 1961 to 1985. The annual AWE for each year is calculated as the average of the AWE for the 12 calendar months. This series uses the 1960 SIC.

3. CANSIM series V252496: This series consists of annual AWE across Canada from 1983 to 2000 with the 1980 SIC.

4. CANSIM series V1796232: This series consists of annual AWE across Canada from 1991 to 2010. This series uses the North American Industry Classification System.

For data between 1961 and 1975, the E49 series and the annual average of the V75249 series produce the same number. Therefore, no adjustment has been made for this period. When there is a change in classification, the impact of the change in classification is smoothed out by taking a weighted average of the two data series.

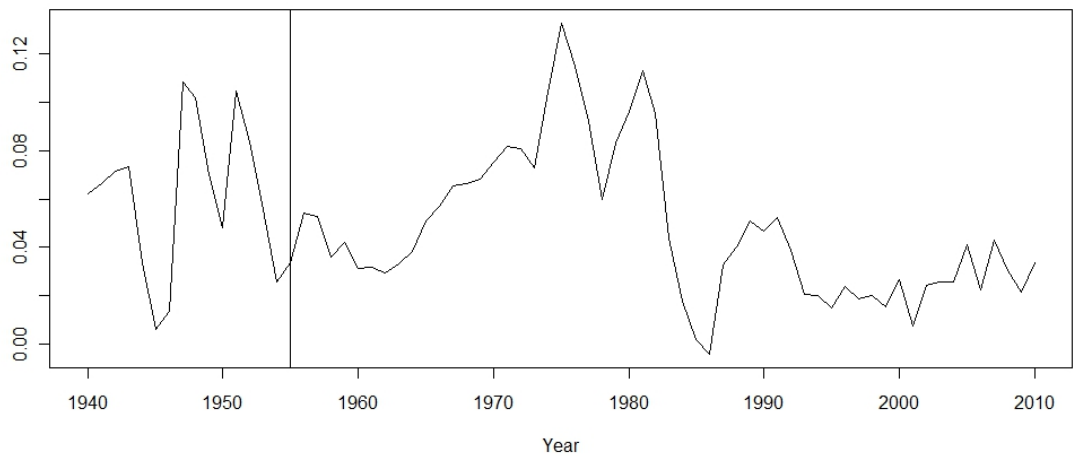

Figure 3.1: Force of Wage Index from 1940 to 2010

The force of annual wage index is calculated as follow:

$$X_W(t) = ln\frac{AWE(t)}{AWE(t-1)}. \tag{3.1.1}$$

In Figure 3.1, the force of annual wage index went through two periods of large volatility: one before 1955 and one around 1980. After 1990, the series has small fluctuations around 2%. To be consistent with the inflation data, only data from 1955 onwards (i.e. right of the vertical line) is used to estimate the parameters.

3.2 Model Selection

3.2.1 AR(1) Model

The ACF and PACF in Figure 3.2 show the pattern of an AR(1) series. As a result, an AR(1) model is appropriate for the force of annual wage index.

3.2.2 ARCH(1) Model

The parameter estimates for the AR(1) model are summarized in Table 3.1. The residual is then calculated based on the parameter estimates. Since there is significant cross-correlation

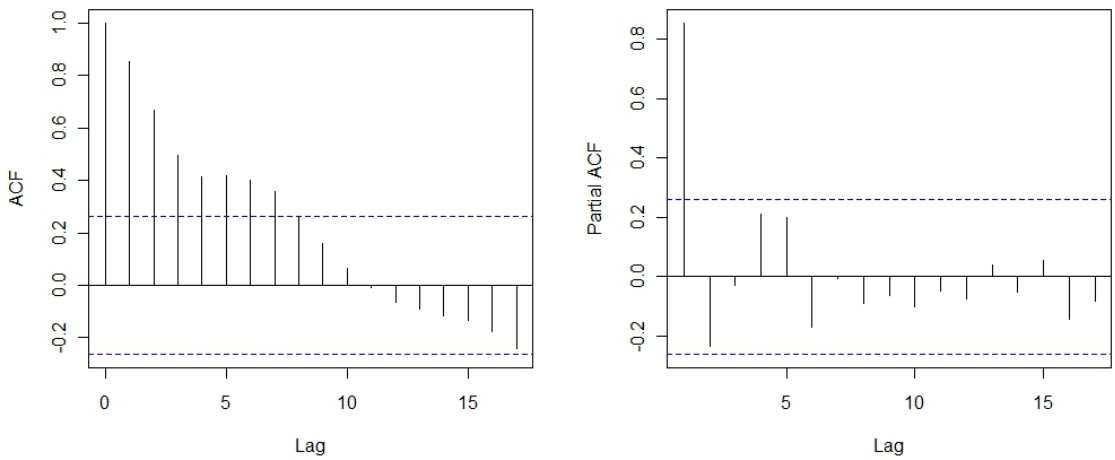

Figure 3.2: Autocorrelation Function and Partial Autocorrelation Function for the Force of Annual Wage Index

at lag minus one between the series and the squared residual, an ARCH(1) model is considered. Note that the cross-correlations at several lags are also significant. However, significant cross-correlations at lags less than minus one imply higher order ARCH models, which is not within the class of models considered. Also, significant cross-correlations at future lags cannot formulate a useful model (as explained in Section 2.2.2). Therefore, the ARCH(F), ARCH(N), and ARCH(P) models are considered.

3.2.3 Transfer Function Models

For the force of annual wage index, transfer function models are also considered with both two lags (TF(2)), and one lag (TF(1)). The force of annual inflation is used as an exogenous variable to explain the force of annual wage index. Figure 3.4 shows the cross-correlation between the force of annual wage index and the force of annual inflation. The cross-correlation is significant at lag zero and minus one, which supports the use of TF(2) and TF(1) models. Note that the cross-correlation is also significant at several other lags; however, significance at lags less than minus one implies a transfer function model with more than two lags, which is not considered. Also, significance at lags larger than zero implies that past observations of the force of annual wage index can be used to explain current observation of the force of

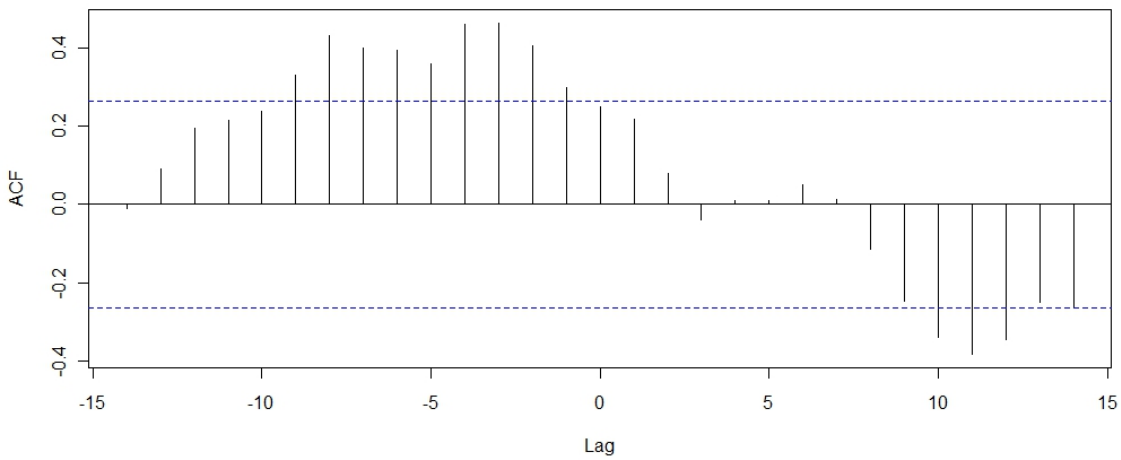

Figure 3.3: Cross-correlation Function for the Force of Annual Wage Index and its Squared AR(1) Residual

annual inflation. This is also not in the class of models considered as it suggests a transfer function model for the force of annual inflation using the force of annual wage index as an exogenous variable.

Moreover, in reality, when the force of annual wage index is very high, employers will want to slow down the wage increase. On the other hand, if the force of annual wage index is very low, the employees will demand bigger wage increase. Therefore, in equilibrium, there is a pressure for the force of annual wage index returning to some base value. In wage negotiation, the force of annual inflation will usually be used as the starting point, and therefore, the force of annual inflation can be considered as some base value. Also, Thury (1979) pointed out that "wage increases are partly a compensation for past or expected future inflation". As a result, the force of inflation can be used to explain part of the increase in wage index. Note that wage negotiation may not occur every year. Also, the timing of the negotiation may be different from the release of the inflation data usually. Therefore, the force of annual inflation in the recent past may affect the current year force of annual wage index. Consequently, TF(2) and TF(1) models are both possible models for the force of annual wage index.

As mentioned in Section 1.2, pension actuaries use the building block approach to come up with the values for future wage increases. Transfer function models can be interpreted

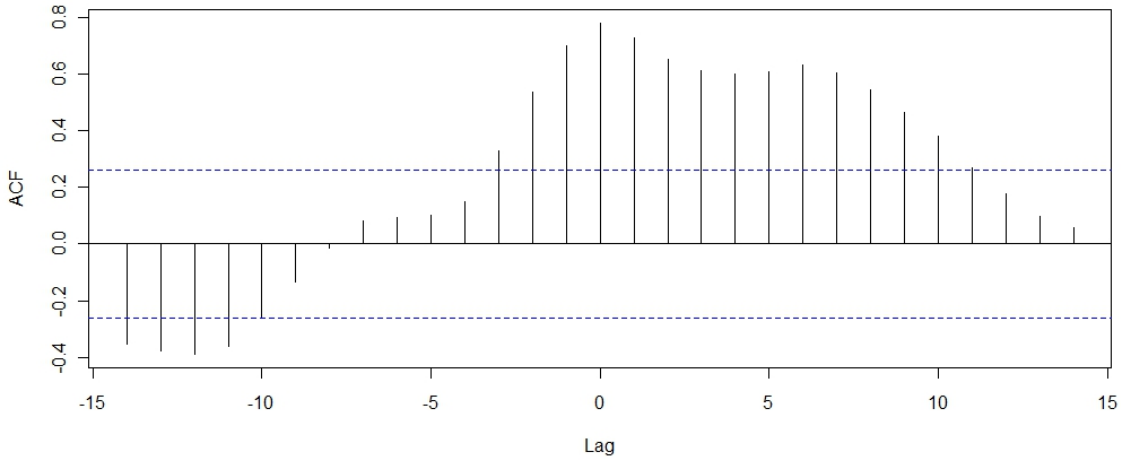

Figure 3.4: Cross-correlation Function for the Force of Annual Wage Index and the Force of Annual Inflation

as a building block where the force of annual wage index is equal to a fraction of the force of annual inflation plus the increase in general productivity and merit modeled as the innovations. The model assumes that the innovations follow an AR(1) model. This is reasonable because the general productivity is correlated with the Gross National Product (GNP) (Özmucur, 2006). Moreover, GNP can be modeled by an autoregressive model (Cochrane, 1994).

3.3 Parameters Estimation

The final parameter estimates, long-term mean, and long-term standard deviation are summarized in Tables 3.1 and 3.2. The asymptotic standard error of the estimation is shown in brackets.

The ARCH(P) model does not produce a stationary standard deviation; therefore, it is not used for the time series projection.

Table 3.1: AR(1)/ARCH(1) Models Parameter Estimates for the Force of Annual Wage Index

Model	AR(1)	ARCH(F)	ARCH(N)	ARCH(P)
$\hat{\mu}_W$	0.0475 (0.0148)	0.0335 (0.0078)	0.0458 (0.0137)	0.0130 (0.0088)
$\hat{\alpha}_W$	0.8586 (0.0691)	0.8978 (0.0862)	0.8519 (0.0784)	0.6970 (0.1949)
$\hat{\sigma}_W$	0.0156 (0.0015)	N/A	N/A	N/A
$\hat{\lambda}_W$	N/A	0.0112 (0.0016)	0.0116 (0.0000)	N/A
$\hat{\gamma}_W$	N/A	0.3501 (0.0890)	0.1820 (0.0413)	1.3297 (0.1268)
LTM	0.0475	0.0335	0.0458	0.0130
LTSD	0.0303	0.0421	0.0291	N/A

Table 3.2: Transfer Function Models Parameter Estimates for the Force of Annual Wage Index

Model	TF(2)	TF(1)
$\hat{\beta}_{W1}$	0.4816 (0.1228)	0.5779 (0.1408)
$\hat{\beta}_{W2}$	0.3458 (0.1215)	N/A
$\hat{\mu}_{We}$	0.0155 (0.0089)	0.0248 (0.0088)
$\hat{\alpha}_{We}$	0.7476 (0.0893)	0.7371 (0.1017)
$\hat{\sigma}_{We}$	0.0127 (0.0012)	0.0136 (0.0013)
LTM	0.0432	0.0441
LTSD	0.0265	0.0243

3.4 Time Series Projection

For the force of annual wage index, two initial values are used: 5% (close to the long-term mean) and 12% (far from the long-term mean). For the transfer function models, RSARCH(P) model is used for the force of annual inflation. Initial values of 3% and 10% (from Section 2.4) for the force of annual inflation are paired with initial values of 5% and 12% for the force of annual wage index respectively.

Figures 3.5, 3.6, 3.7, and 3.8 show the mean, standard deviation, 5th and 95th percentiles, and the minimum and maximum of the projection data, respectively, for the AR(1) and ARCH(1) models. Figures 3.9, 3.10, 3.11, and 3.12 show the mean, standard deviation, 5th and 95th percentiles, and the minimum and maximum of the projection data, respectively, for the transfer function models.

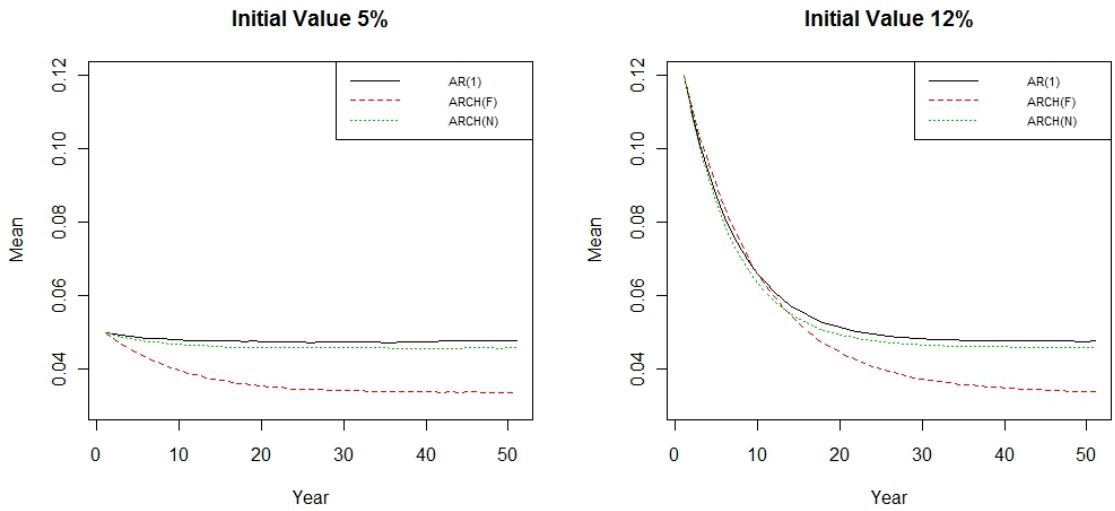

Figure 3.5: AR(1)/ARCH(1) Models Mean of Projection for the Force of Annual Wage Index

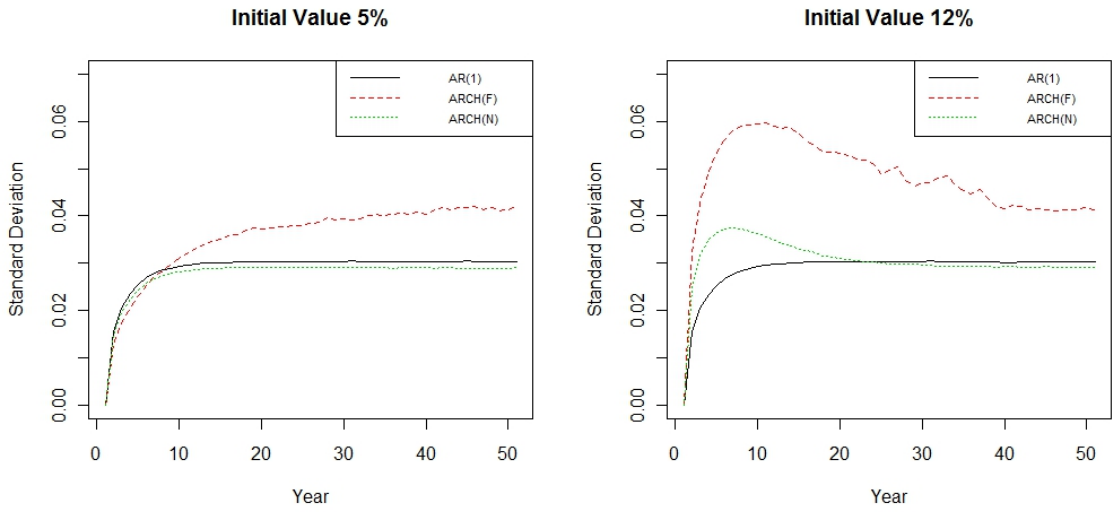

Figure 3.6: AR(1)/ARCH(1) Models Standard Deviation of Projection for the Force of Annual Wage Index

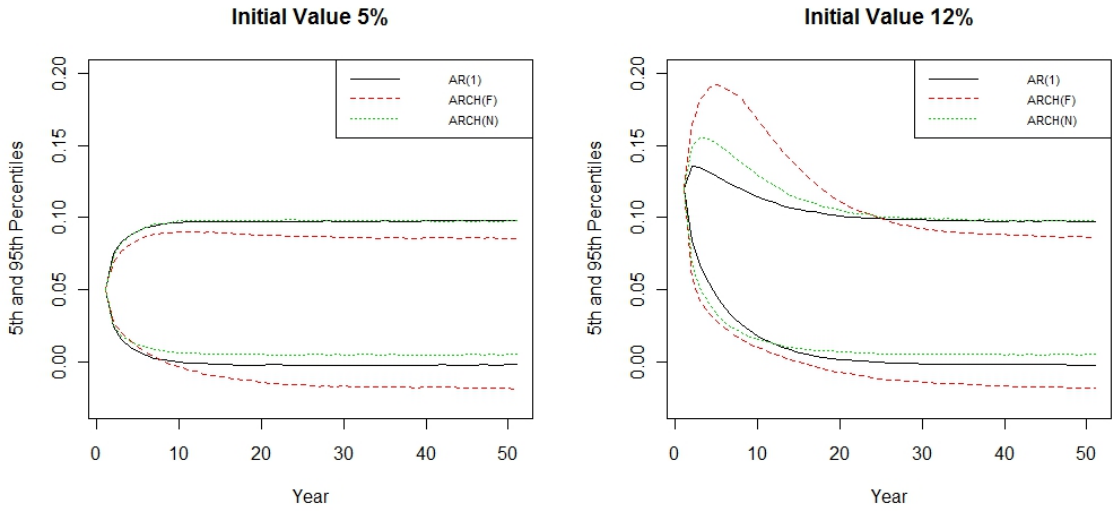

Figure 3.7: AR(1)/ARCH(1) Models 5th and 95th Percentiles of Projection for the Force of Annual Wage Index

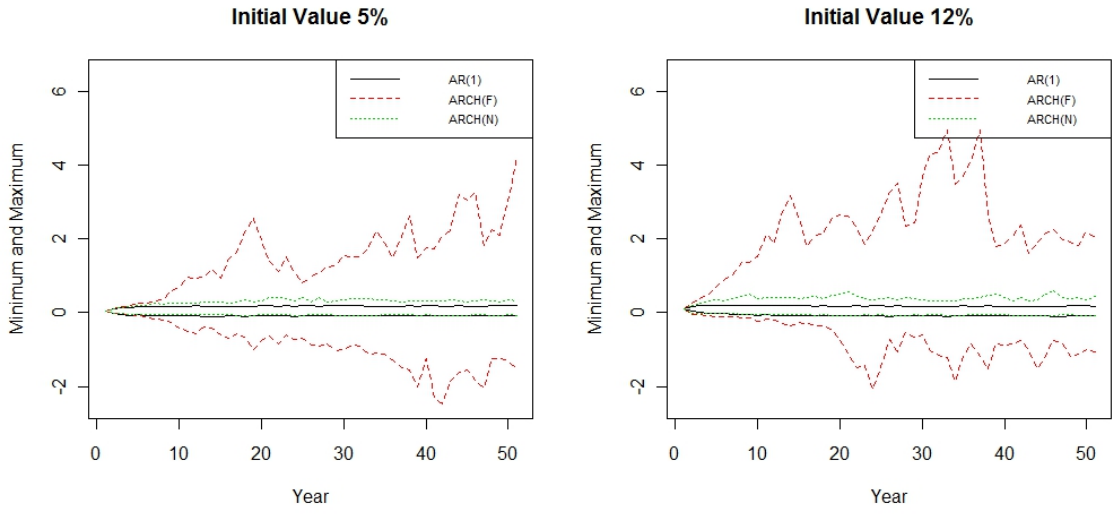

Figure 3.8: AR(1)/ARCH(1) Models Minimum and Maximum of Projection for the Force of Annual Wage Index

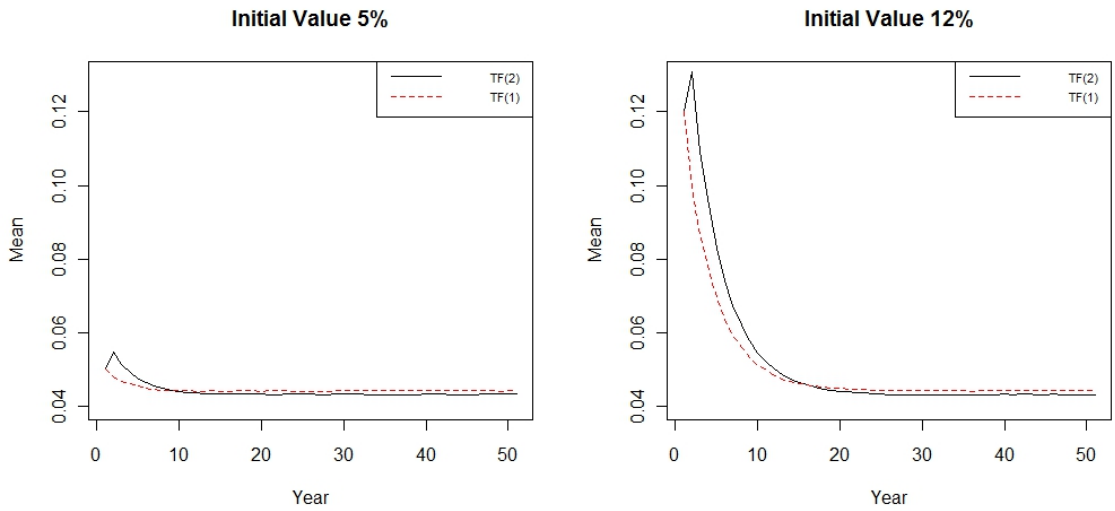

Figure 3.9: Transfer Function Models Mean of Projection for the Force of Annual Wage Index

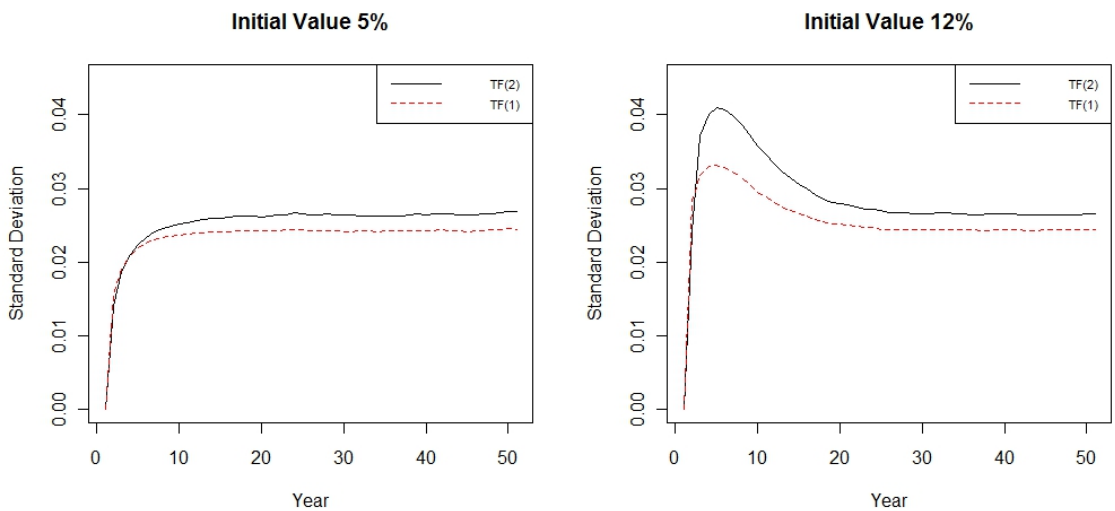

Figure 3.10: Transfer Function Models Standard Deviation of Projection for the Force of Annual Wage Index

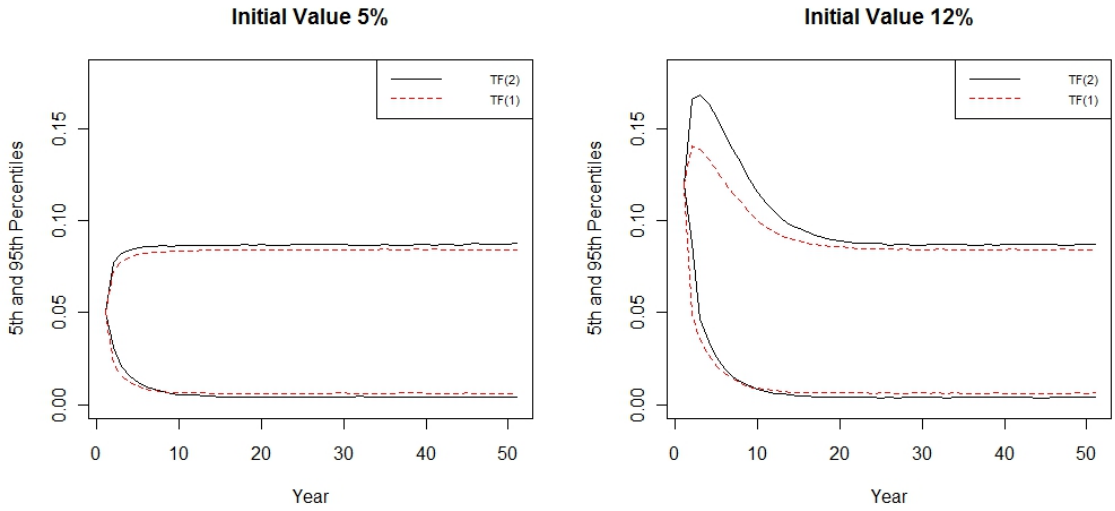

Figure 3.11: Transfer Function Models 5th and 95th Percentiles of Projection for the Force of Annual Wage Index

Figure 3.12: Transfer Function Models Minimum and Maximum of Projection for the Force of Annual Wage Index

3.5 Final Model

The long-term mean for the ARCH(F) model is significantly lower than the AR(1) and ARCH(N) models. This may be a problem for pension plan projections because one of the risks faced by employers in DB pension plans and employees in DC pension plans is wage increases. For DB pension plan employers, in general, increased wages mean more benefits, and therefore, costs and contributions. For DC pension plan employees, increased wages mean more employer contributions but also more retirement income required to maintain the pre-retirement standard of living. Therefore, using an ARCH(F) model for the force of annual wage index may underestimate the risk by having a lower long-term mean force of annual wage index.

In Figure 3.6, the standard deviation for the ARCH(F) model seems to be just about to converge to the long-term value after 50 years. This is driven by the large extreme values shown in Figure 3.8. Combined with the lower mean mentioned above, the ARCH(F) model is not appropriate to model the force of annual wage index for pension plan projections.

The dynamics of the AR(1) and ARCH(N) models are somewhat similar as shown in Figures 3.5-3.8, except for the fact that the ARCH(N) model has larger variance before stationarity when the initial value is away from the mean. The ARCH(N) model also produces slightly larger range of projected values than the AR(1) model. Again, all else being equal, models with conditional heteroscedasticity are preferable to an AR(1) model for pension plan projections.

For TF(2) and TF(1) models, the dynamics are very similar, with the exception that the standard deviation for TF(2) is larger than TF(1). This is due to the underlying force of annual inflation being modeled by the RSARCH(P) model. Since $\hat{\beta}_{W2}$ is larger than twice of its own standard error, β_{W2} may be significant and should not be eliminated. Therefore, the TF(2) model would be more appropriate than the TF(1) model.

Akaike (1974) proposed the Akaike Information Criterion (AIC) to rank competing models. AIC is defined as

$$AIC = 2k - 2L \tag{3.5.1}$$

where k is the number of parameters in the model and L is the maximum log-likelihood value. The model with the lowest AIC is the most appropriate model among the competing models. The AIC for the TF(2) and TF(1) models are -314.32 and -308.74 respectively.

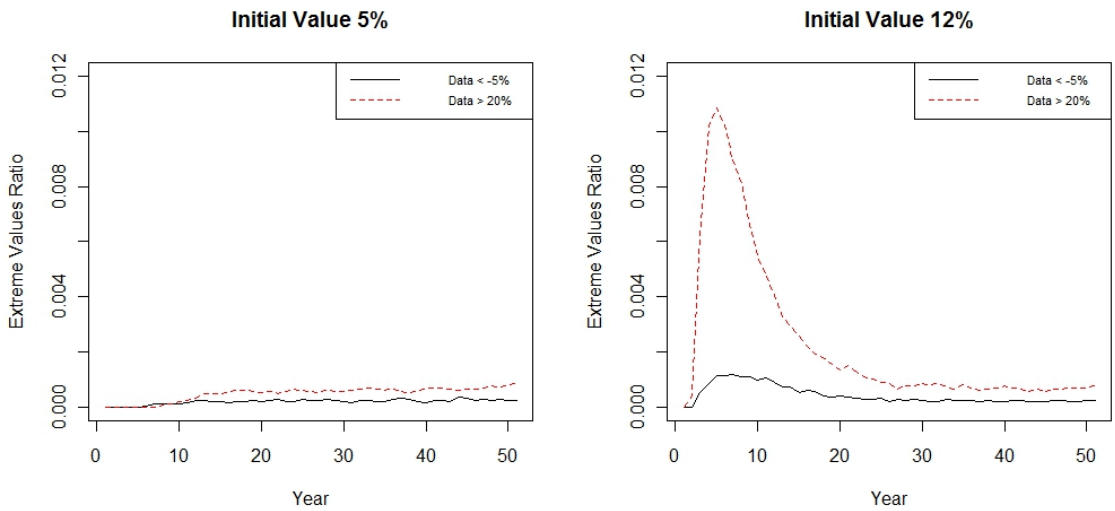

Figure 3.13: Extreme Value Ratios of Projection for the Force of Annual Wage Index TF(2) Model

This also suggest that the TF(2) model is more appropriate than the TF(1) model.

Figure 3.13 shows the extreme value ratios for the TF(2) model. Extreme value thresholds are defined the same as Section 2.5. The shape of the curve is similar to Figure 2.8 because it is based on the same simulated force of annual inflation. The frequency of extreme values occurrence is somewhat less than the force of annual inflation; therefore, it is also acceptable for pension plan projections.

In conclusion, the TF(2) model is preferable to all the other models because it is justifiable in a statistical sense. At the same time, this model incorporates the force of annual inflation, and as a result, the regime switching model chosen for the force of annual inflation.

Chapter 4

Long-term Interest Rate

4.1 Data

The monthly average yield for Government of Canada marketable bonds with maturity over 10 years is available at Statistics Canada from 1936 to 2010 based on CANSIM series V122487. The average yield for the last Wednesday of December of each year is used as a proxy for the annual long-term interest rate.

In Figure 4.1, the annual long-term interest rate was somewhat flat before 1955. It then increased and reached a maximum around 1980. Since then, the annual long-term interest rate has been declining. Excluding the pre-1955 data introduces more volatility into the data, and thus, into the models implied by the estimated parameters. However, given the recently increased possibility of government bonds defaulting in the United States and in Europe, more volatility in the Canadian economy can be expected (Simms, 2011). As a result, it is reasonable to assume similar volatility will continue into the future for the annual long-term interest rate; and therefore, only data from 1955 onwards (i.e. right of the vertical line) is used to estimate the parameters.

4.2 Model Selection

4.2.1 AR(1) Model

The ACF and PACF in Figure 4.2 show the pattern of an AR(1) series. As a result, an AR(1) model is appropriate for the annual long-term interest rate.

Figure 4.1: Annual Long-term Interest Rate from 1936 to 2010

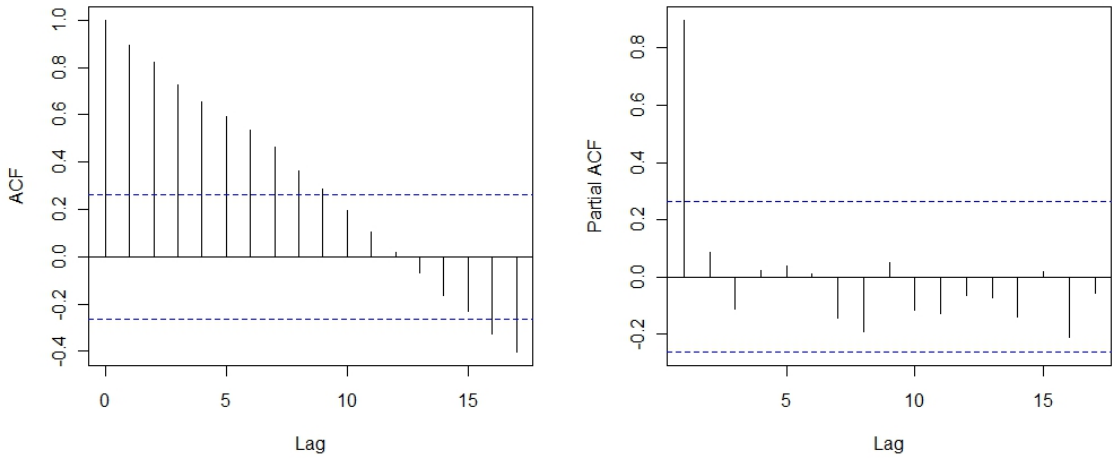

Figure 4.2: Autocorrelation Function and Partial Autocorrelation Function for the Annual Long-term Interest Rate

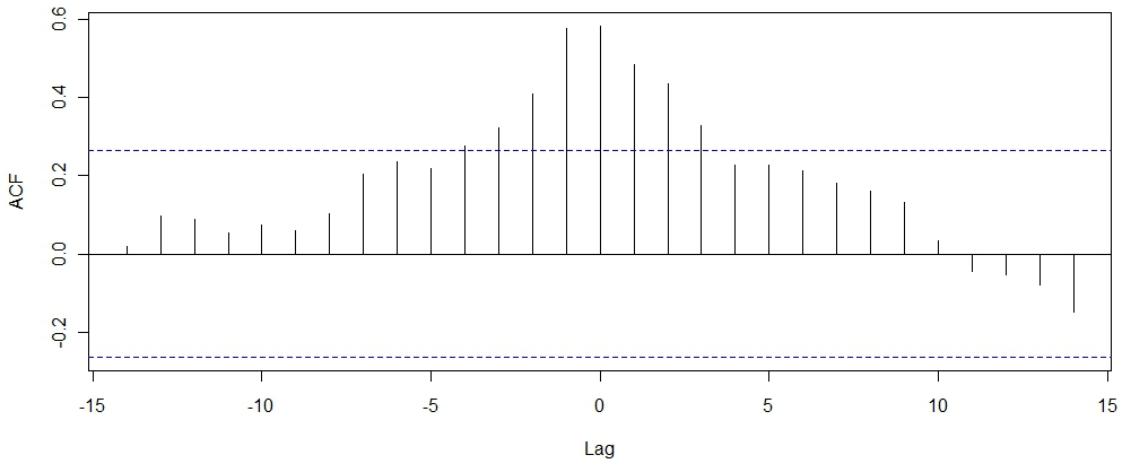

Figure 4.3: Cross-correlation Function for the Annual Long-term Interest Rate and its Squared AR(1) Residual

4.2.2 ARCH(1) Model

The parameter estimates for the AR(1) model are summarized in Table 4.1. The residual is then calculated based on the parameter estimates. Since there is significant cross-correlation at lag minus one between the series and the squared residual in Figure 4.3, an ARCH(1) model is considered. Again, significance in cross-correlations at other lags is either not considered or not useful. Therefore, the ARCH(F), ARCH(N), and ARCH(P) models are considered.

4.2.3 Transfer Function Models

For annual long-term interest rate, transfer function models are also considered with both two lags and one lag. The force of annual inflation is used as an exogenous variable to explain the annual long-term interest rate. Figure 4.4 shows the cross-correlation between the annual long-term interest rate and the force of annual inflation. The cross-correlation is significant at lag zero and minus one, which supports the use of the TF(2) and TF(1) models. As in Section 4.2.2, significance of the cross-correlation at other lags is not considered.

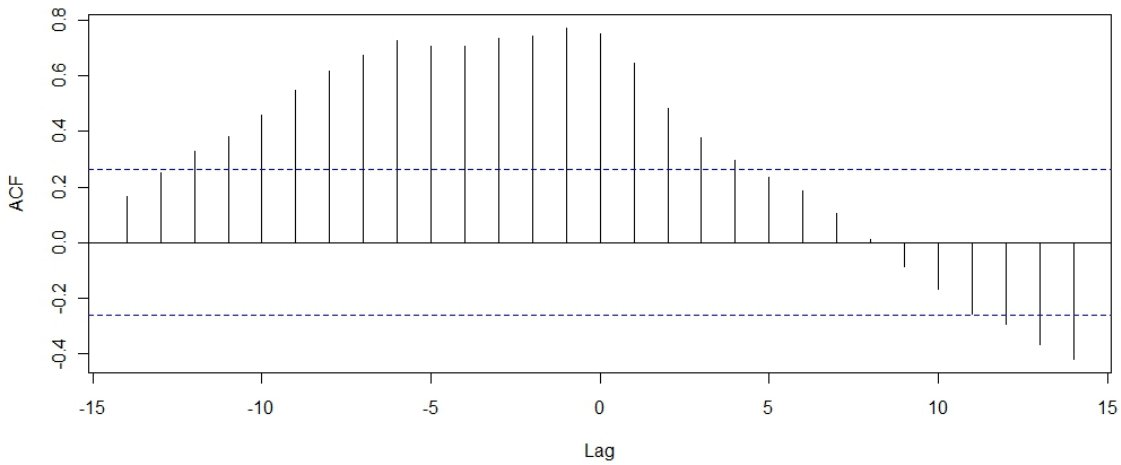

Figure 4.4: Cross-correlation Function for the Annual Long-term Interest Rate and the Force of Annual Inflation

4.3 Parameters Estimation

The final parameter estimates, long-term mean, and long-term standard deviation are summarized in Tables 4.1 and 4.2. The asymptotic standard error of the estimation is shown in brackets.

Table 4.1: AR(1)/ARCH(1) Models Parameter Estimates for the Annual Long-term Interest Rate

Model	AR(1)	ARCH(F)	ARCH(N)	ARCH(P)
$\hat{\mu}_L$	0.0730 (0.0186)	0.0421 (0.0134)	0.0739 (0.0256)	0.0739 (0.0256)
$\hat{\alpha}_L$	0.9269 (0.0502)	0.9976 (0.0429)	0.9488 (0.0483)	0.9488 (0.0483)
$\hat{\sigma}_L$	0.0101 (0.0010)	N/A	N/A	N/A
$\hat{\lambda}_L$	N/A	0.0053 (0.0010)	0.0000 (0.0000)	N/A
$\hat{\gamma}_L$	N/A	0.2113 (0.0705)	0.1230 (0.0117)	0.1230 (0.0117)
LTM	0.0730	0.0421	0.0739	0.0739
LTSD	0.0269	N/A	0.0312	0.0312

The ARCH(F) model does not produce a stationary standard deviation, and the ARCH(N) model is essentially the same as the ARCH(P) model. Therefore, neither model is used for the time series projection.

Table 4.2: Transfer Function Models Parameter Estimates for the Annual Long-term Interest Rate

Model	TF(2)	TF(1)
$\hat{\beta}_{L1}$	0.1269 (0.0951)	0.1327 (0.0966)
$\hat{\beta}_{L2}$	0.1728 (0.0970)	N/A
$\hat{\mu}_{Le}$	0.0613 (0.0129)	0.0674 (0.0162)
$\hat{\alpha}_{Le}$	0.8893 (0.0668)	0.9147 (0.0563)
$\hat{\sigma}_{Le}$	0.0096 (0.0009)	0.0099 (0.0009)
LTM	0.0713	0.0719
LTSD	0.0221	0.0248

4.4 Time Series Projection

For the annual long-term interest rate, two initial values are used: 7% (close to the long-term mean) and 14% (far from the long-term mean). For the transfer function models, the RSARCH(P) model is used for the force of annual inflation. Initial values of 3% and 10% (from Section 2.4) for the force of annual inflation are paired with initial values of 7% and 14% for the annual long-term interest rate respectively.

Figures 4.5, 4,6, 4.7, and 4.8 show the mean, standard deviation, 5th and 95th percentiles, and the minimum and maximum of the projection data, respectively, for the AR(1) and ARCH(1) models. Figures 4.9, 4.10, 4.11, and 4.12 show the mean, standard deviation, 5th and 95th percentiles, and the minimum and maximum of the projection data, respectively, for the transfer function models.

4.5 Final Model

The AR(1) and ARCH(P) models have similar dynamics as illustrated in Figures 4.5-4.8 with some exceptions. The ARCH(P) model has larger standard deviation than the AR(1) model regardless of the initial value. Both models have high autocorrelation factor $\hat{\alpha}_L$, and this high value produces a unique feature for the ARCH(P) model. The minimum values simulated using the ARCH(P) model do not go below zero as seen in Figure 4.8. Negative interest rates rarely occur in reality because interest rates are monitored by the Government ("Key Interest Rate", n.d.) and negative interest rates are not good for the economy. Also, for pension plan projections, negative interest rates imply setting aside a larger amount now in order to pay a smaller amount in the future. This goes against logic because, at

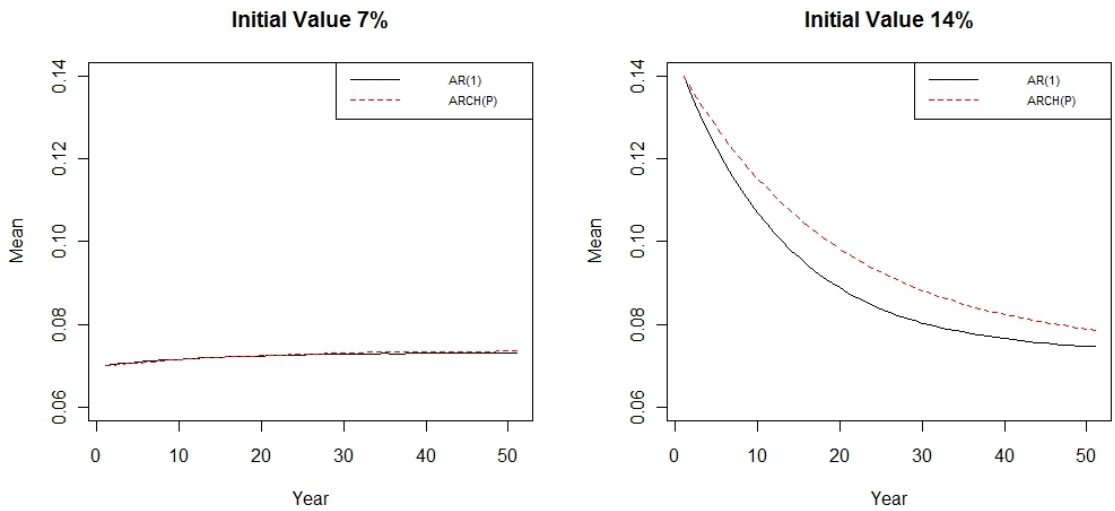

Figure 4.5: AR(1)/ARCH(1) Models Mean of Projection for the Annual Long-term Interest Rate

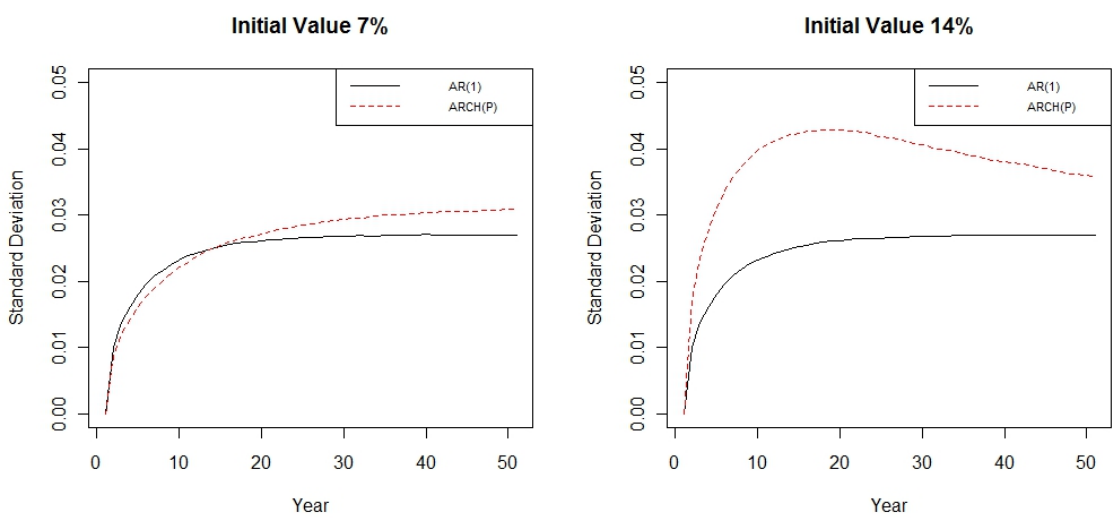

Figure 4.6: AR(1)/ARCH(1) Models Standard Deviation of Projection for the Annual Long-term Interest Rate

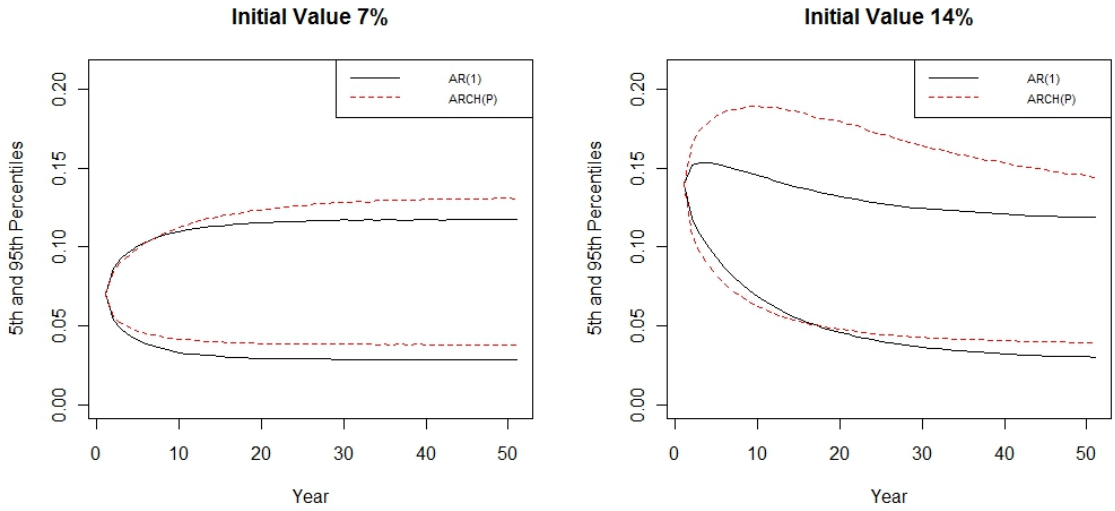

Figure 4.7: AR(1)/ARCH(1) Models 5th and 95th Percentiles of Projection for the Annual Long-term Interest Rate

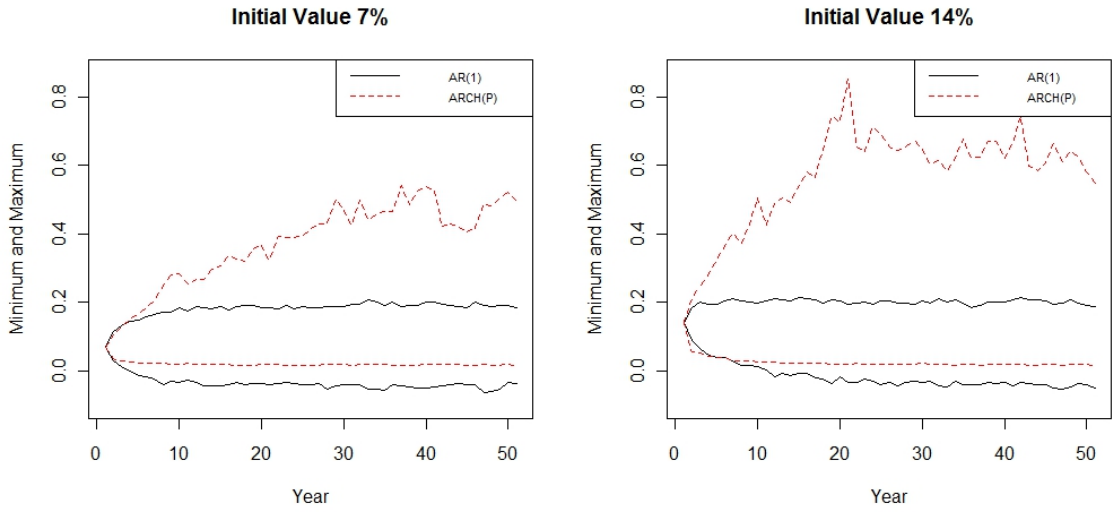

Figure 4.8: AR(1)/ARCH(1) Models Minimum and Maximum of Projection for the Annual Long-term Interest Rate

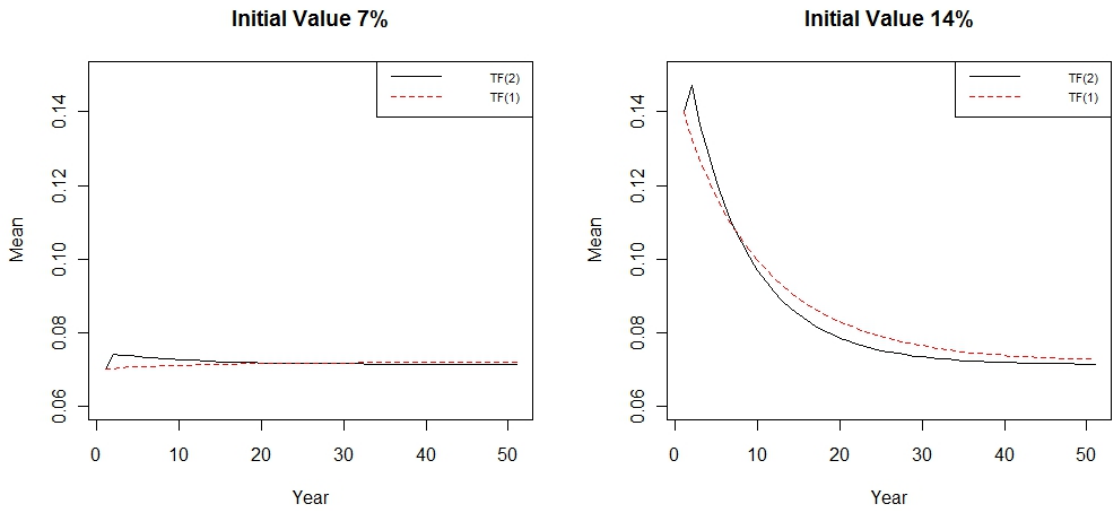

Figure 4.9: Transfer Function Models Mean of Projection for the Annual Long-term Interest Rate

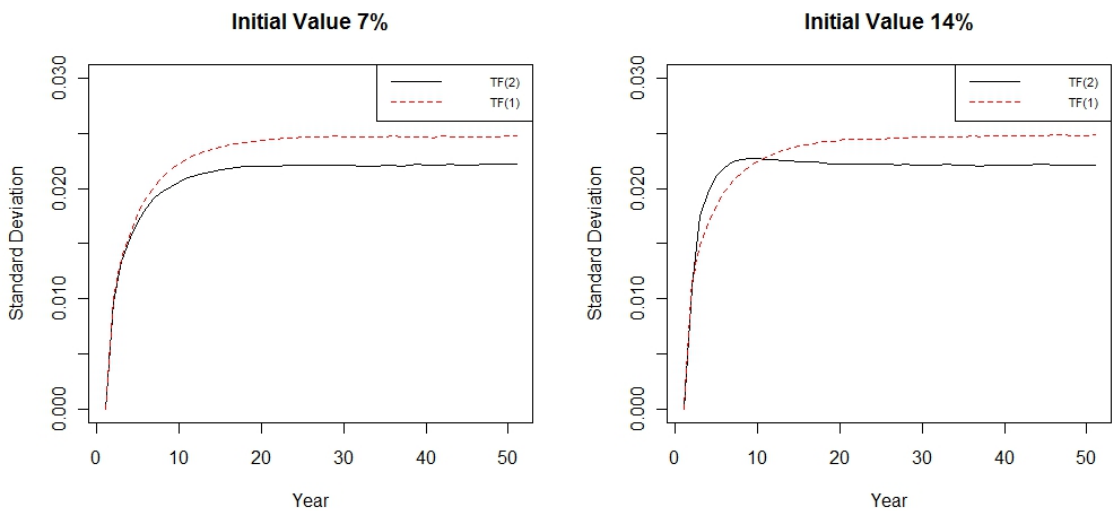

Figure 4.10: Transfer Function Models Standard Deviation of Projection for the Annual Long-term Interest Rate

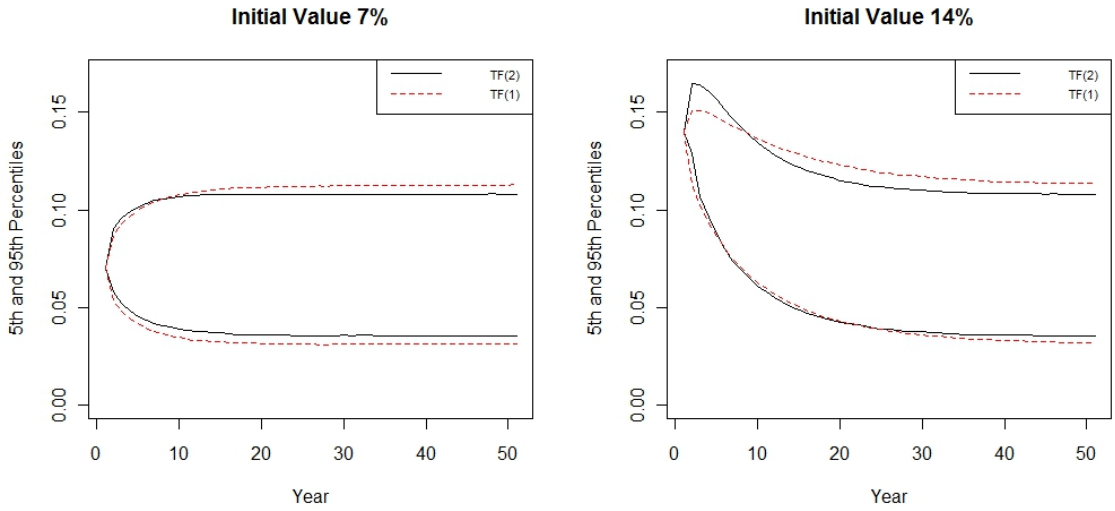

Figure 4.11: Transfer Function Models 5th and 95th Percentiles of Projection for the Annual Long-term Interest Rate

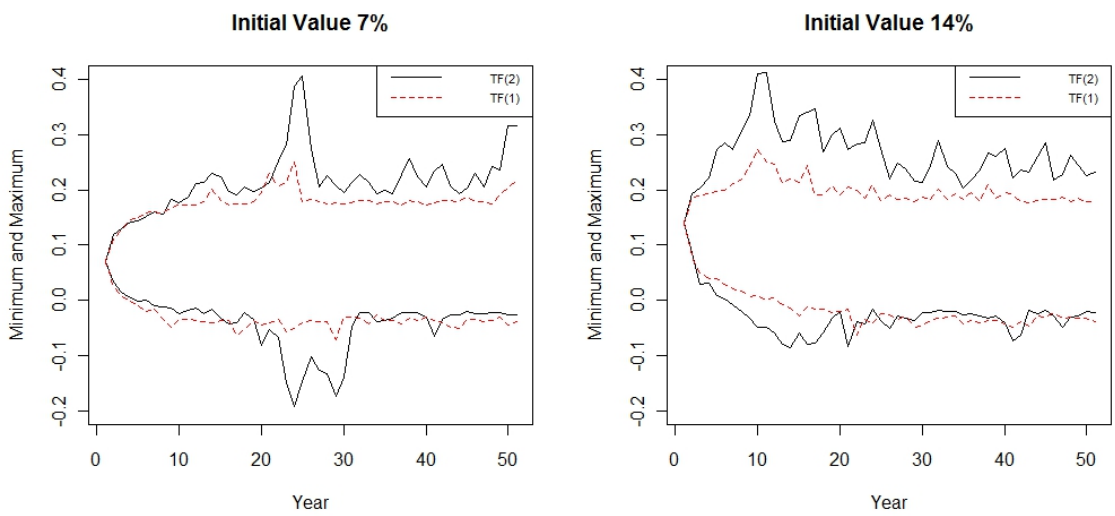

Figure 4.12: Transfer Function Models Minimum and Maximum of Projection for the Annual Long-term Interest Rate

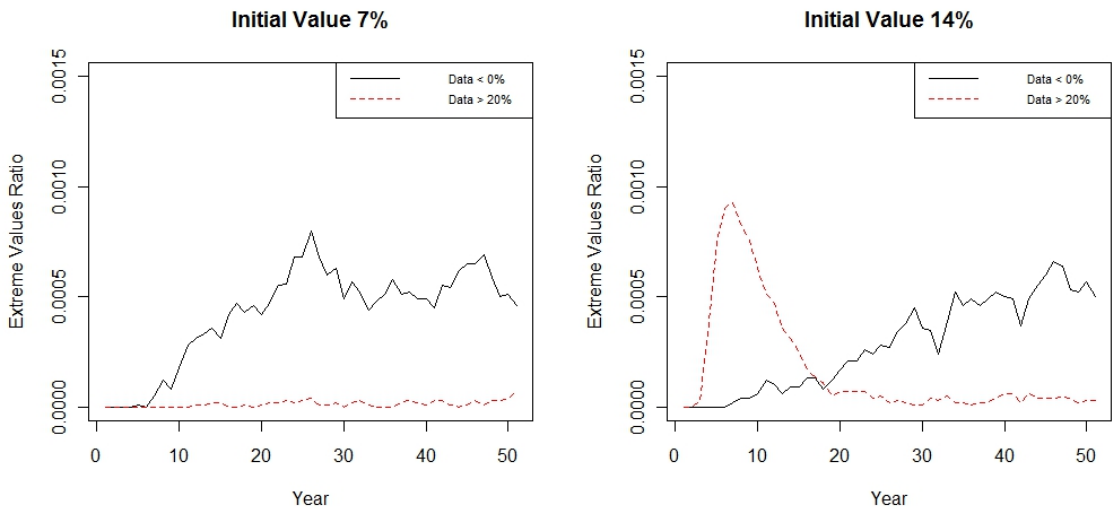

Figure 4.13: Extreme Value Ratios of Projection for the Annual Long-term Interest Rate TF(2) Model

the least, setting aside an amount equal to the amount to be paid in the future should be enough. Therefore, negative interest rates may not produce meaningful results for pension plan projections.

Neither $\hat{\beta}_{L1}$ nor $\hat{\beta}_{L2}$ is greater than twice of its own standard error. However, this does not necessarily imply the transfer function models are not appropriate. First, Figure 4.4 shows that the annual long-term interest rate and the force of annual inflation have significant cross-correlations at different lags. Second, the parameters may not be normally distributed because both $\hat{\beta}_{L1}$ and $\hat{\beta}_{L2}$ should be positive as indicated by the positive correlation in Figure 4.4. Also, standard errors calculated using Fisher information matrix assume the parameters have asymptotic normal distribution. Since only 56 observations are used to estimate the parameters, the asymptotic behaviour may not be achieved. Third, the correlation between the annual long-term interest rate and the force of annual inflation is a combination of $\hat{\beta}_{L1}$, $\hat{\beta}_{L2}$, and $\hat{\alpha}_{Le}$. Since $\hat{\alpha}_{Le}$ is very large, it is reasonable to have both $\hat{\beta}_{L1}$ and $\hat{\beta}_{L2}$ relatively small so that the cross-correlation implied by the parameters estimates is consistent with the data. In this report, transfer function models are considered to be reasonable models for the annual long-term interest rates.

Figures 4.9-4.12 show the comparison between the TF(2) and TF(1) models. The most significant difference between the two models is that, while the TF(1) model has a slightly

higher long-term standard deviation, the TF(2) model has generally larger range than the TF(1) model and is more sensitive to big changes in the underlying force of annual inflation. The AIC for the TF(2) and TF(1) models are -344.45 and -343.28 respectively. Therefore, the TF(2) model is again preferable to the TF(1) model.

Note that the TF(2) model does produce negative interest rates; therefore, the investigation of the extreme values is important for annual long-term interest rate. In Figure 4.13, extreme value thresholds are defined as annual long-term interest rate greater than 20% or smaller than 0%. The extreme value thresholds are tighter than those in Sections 2.5 and 3.5 because the extent of negative interest rates needs to be understood. The frequency of occurrence of extreme values is very low as illustrated in Figure 4.13. In conclusion, the TF(2) model is the preferred model for annual long-term interest rate.

Chapter 5

Equity Return

5.1 Data

For annual Canadian equity return (CER), the monthly S&P/TSX composite index (formerly TSE 300 index) is available at Statistics Canada from 1952 to 2010 based on CANSIM series V7668. This index includes the equity prices of the largest companies on the Toronto Stock Exchange. The Canadian equity index (CEI) in December for each year, t, is used to calculate the force of annual CER, X_C, using the following equation:

$$X_C(t) = ln\frac{CEI(t)}{CEI(t-1)}. \tag{5.1.1}$$

Figure 5.1 shows a plot of the force of annual CER.

For annual global equity return (GER), the annual MSCI World index, as a proxy for the global equity index (GEI), is available at the website of Morgan Stanley Capital International (MSCI) from 1969 to 2010. This index includes 24 developed market country indices. The force of annual GER, X_G, is calculated using the following equation:

$$X_G(t) = ln\frac{GEI(t)}{GEI(t-1)}. \tag{5.1.2}$$

Figure 5.2 shows a plot of the force of annual GER.

Figure 5.1: Force of Annual Canadian Equity Return from 1953 to 2010

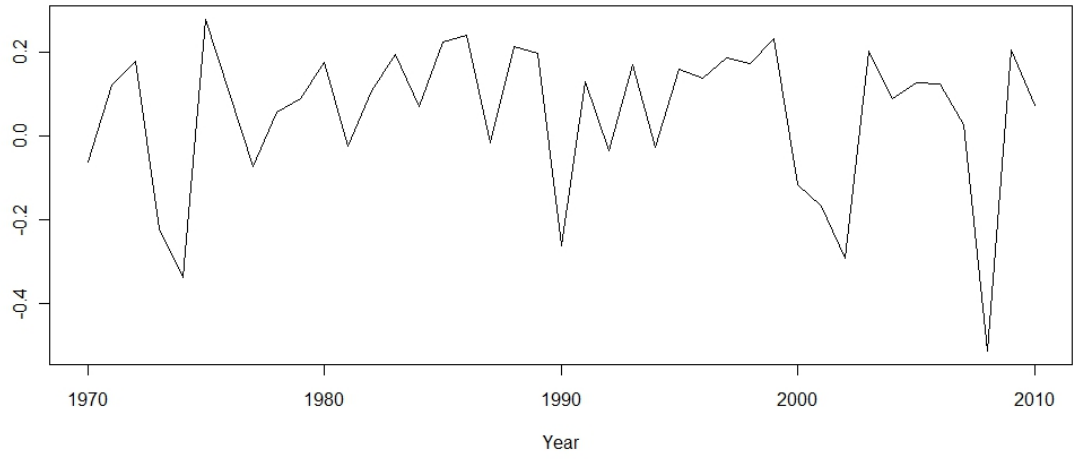

Figure 5.2: Force of Annual Global Equity Return from 1970 to 2010

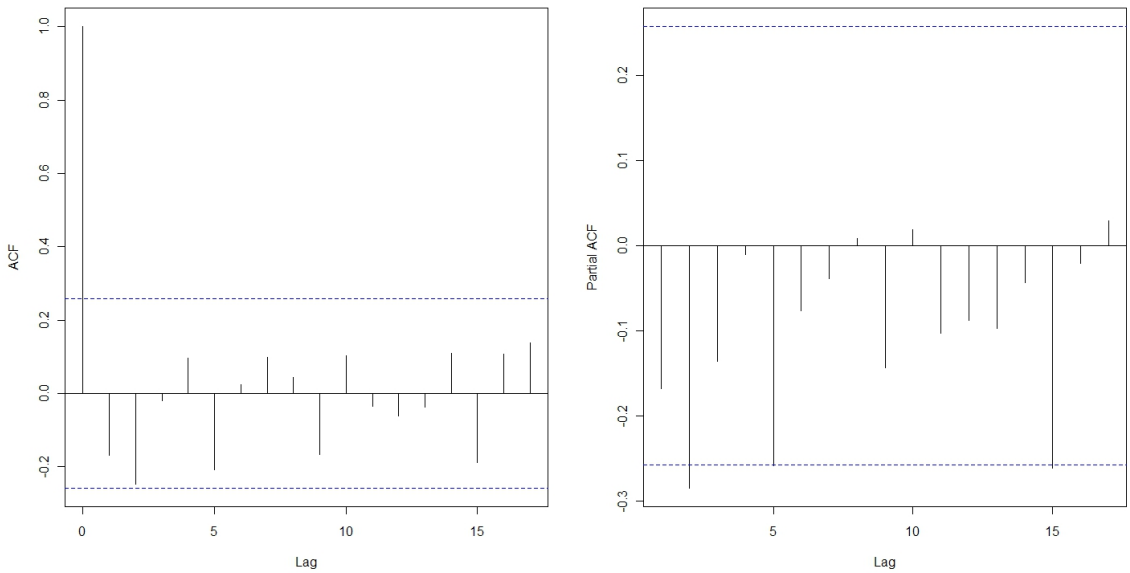

Figure 5.3: Autocorrelation Function and Partial Autocorrelation Function for the Force of Annual Canadian Equity Return

5.2 Model Selection

5.2.1 White Noise Model

The ACF and PACF in Figures 5.3 and 5.4 both show the pattern of a white noise series. As a result, a white noise (WN) model is appropriate for both the forces of annual CER and GER.

5.2.2 ARCH(1) Model

The parameter estimates for the WN model are summarized in Table 5.1. The residual is then calculated based on the parameter estimates. Figures 5.5 and 5.6 show the cross-correlation of the series with its squared residual. Since there is no significant cross-correlation other than at lag zero for both series, ARCH models are not considered.

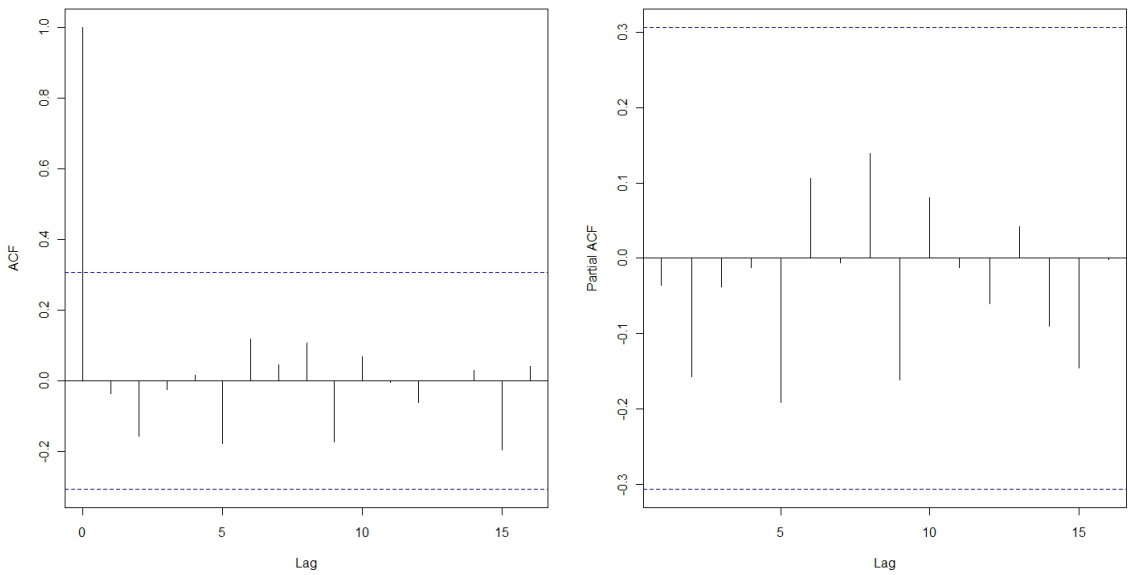

Figure 5.4: Autocorrelation Function and Partial Autocorrelation Function for the Force of Annual Global Equity Return

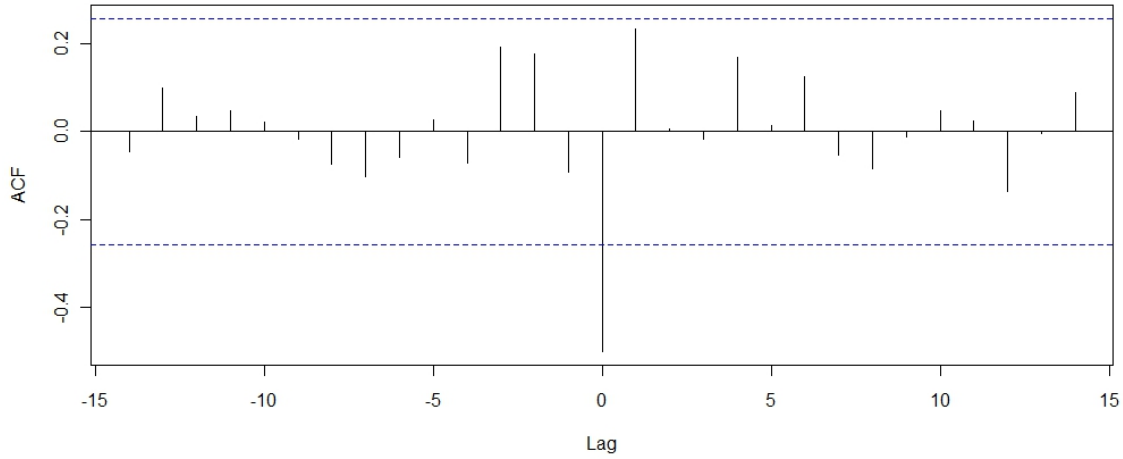

Figure 5.5: Cross-correlation Function for the Force of Annual Canadian Equity Return and its Squared WN Residual

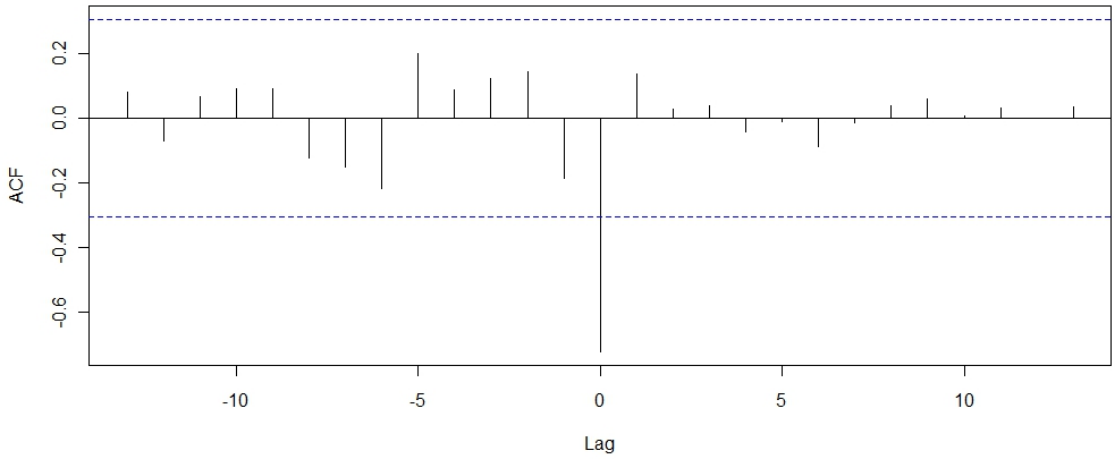

Figure 5.6: Cross-correlation Function for the Force of Annual Global Equity Return and its Squared WN Residual

5.2.3 Regime Switching Vector White Noise Model

For both of the forces of annual CER and GER, regime switching models are considered. The regime for the forces of annual equity return is assumed to be independent of the regime for the force of annual inflation (see Section 2.2.3). On the other hand, the regime for both of the forces of annual CER and GER are assumed to be the same. Therefore, a regime switching vector white noise (RSWN) model is considered. Figure 5.7 shows the cross-correlation between the two series being significant at lag zero, which supports the idea of a vector regime switching model.

Note that, for this model, the only correlation between the force of annual CER and the force of annual GER is through the regime. That is, the innovations of the two series are assumed to be independent. Also, in order to formulate a vector model, the same number of observations are required for both series. Therefore, only data from 1970 to 2010 is used for both series to estimate the parameters of the RSWN model.

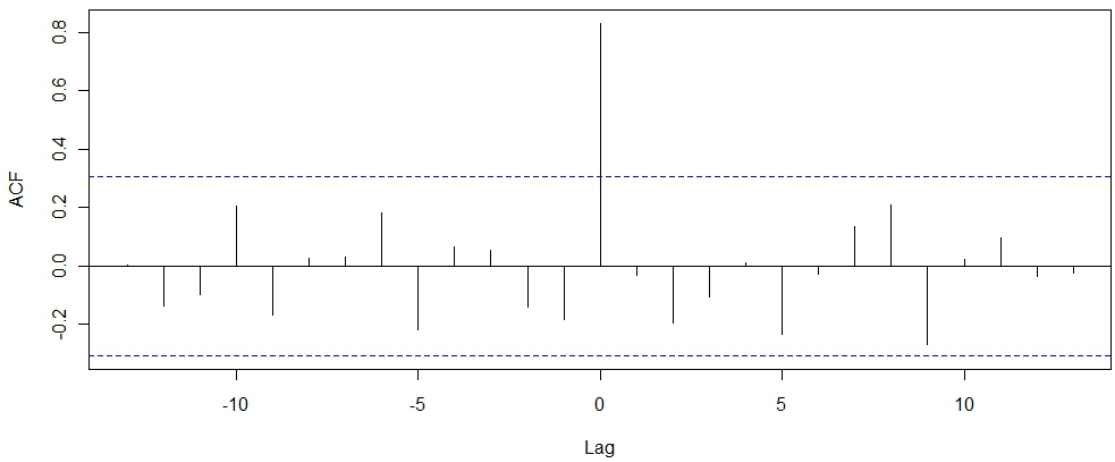

Figure 5.7: Cross-correlation Function for the Force of Annual Canadian Equity Return and the Force of Annual Global Equity Return

5.3 Parameters Estimation

The final parameter estimates are summarized in Tables 5.1 and 5.2 along with the asymptotic standard error of the estimation (shown in brackets), the long-term mean, and the long-term standard deviation.

Table 5.1: WN Model Parameter Estimates for the Forces of Annual CER and GER

Series	CER	GER
$\hat{\mu}$	0.0625 (0.0211)	0.0535 (0.0277)
$\hat{\sigma}$	0.1609 (0.0149)	0.1771 (0.0196)
LTM	0.0625	0.0535
LTSD	0.1609	0.1771

5.4 Time Series Projection

For the forces of annual CER and GER, the last observation is used as the initial value because a white noise model is used, and therefore, the initial value does not change the dynamics of the model. Figures 5.8-5.11 show the mean, standard deviation, 5th and 95th percentiles, the minimum and maximum, and the distribution of the accumulated rates of

Table 5.2: RSWN Model Parameter Estimates for the the Forces of Annual CER and GER

Series	CER	GER
$\hat{\mu}_1$	-0.0991	-0.1329
$\hat{\sigma}_1$	0.1420	0.1618
$\hat{\mu}_2$	0.1545	0.1589
$\hat{\sigma}_2$	0.0893	0.0597
$\hat{\rho}_{12}$	0.7385	
$\hat{\rho}_{21}$	0.4155	
LTM	0.0632	0.0538
LTSD	0.1649	0.1772

return at year 50 of the projection data for the WN and RSWN models.

5.5 Final Model

For both forces of annual CER and GER, the WN and RSWN models have essentially the same long-term mean and standard deviation. The slight difference in standard deviation for the force of annual CER is due to using different data to estimate the parameters of the WN and RSWN models.

Hardy (2003) pointed out that the density function for the RSWN model has a fatter left tail than the one for the WN model in the long term. The 5th and 95th percentiles correspond to a smaller return for the RSWN model than the WN model as shown in Figure 5.10. Also, in Figure 5.11, although the minimum value is not that different between the two models, the maximum value is clearly shifted downwards for the RSWN model. Moreover, Figure 5.12 suggests that the probability of the accumulated rates of return at year 50 being less than 5 is greater for the RSWN model. All three figures confirmed Hardy's theory.

Fatter left tail is an important feature for the RSWN model in pension plan projections because it highlights the risk faced by both employers and employees. In a low return environment, employers in DB pension plans are more likely to contribute large amounts, which increases their financial risk. At the same time, employees in DC pension plans are more likely to accumulate insufficient assets to retire and maintain a similar standard of living after retirement. The RSWN model demonstrates that the path of the rate of return can affect the accumulated rates of return significantly, even though the mean and standard deviation of the individual annual rate of return are the same as the WN model every year.

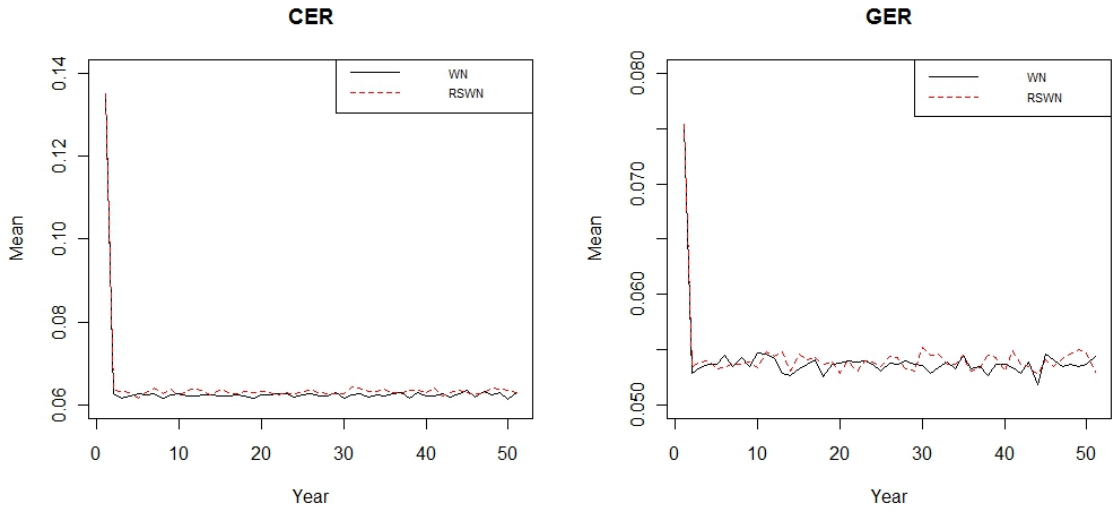

Figure 5.8: WN/RSWN Models Mean of Projection for the Forces of Annual Canadian Equity Return and Global Equity Return

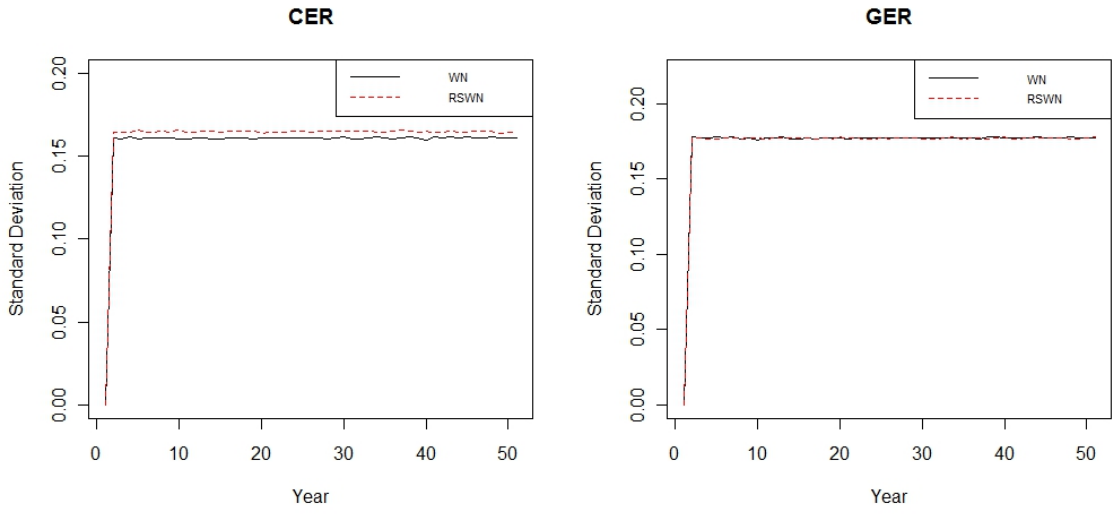

Figure 5.9: WN/RSWN Models Standard Deviation of Projection for the Forces of Annual Canadian Equity Return and Global Equity Return

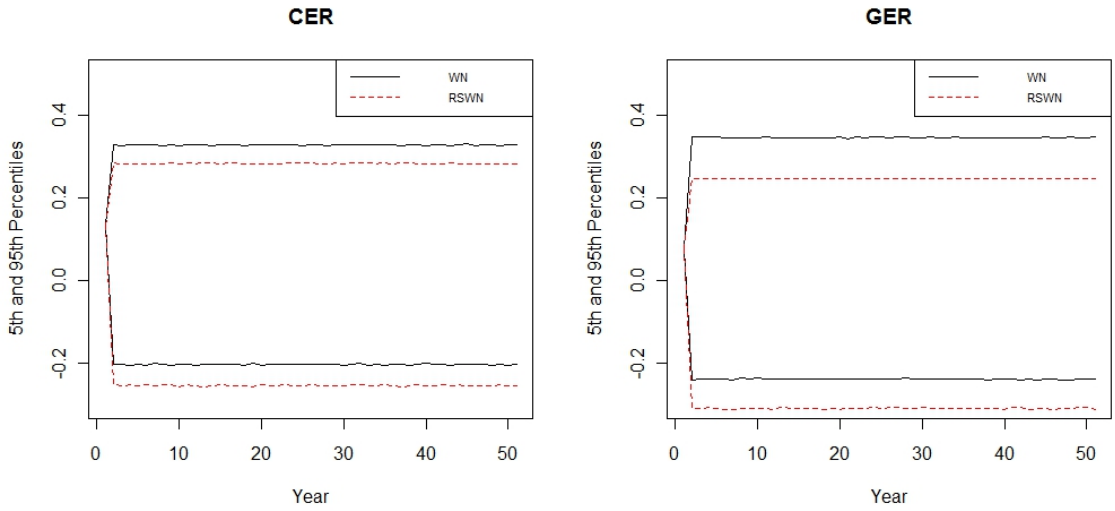

Figure 5.10: WN/RSWN Models 5th and 95th Percentiles of Projection for the Forces of Annual Canadian Equity Return and Global Equity Return

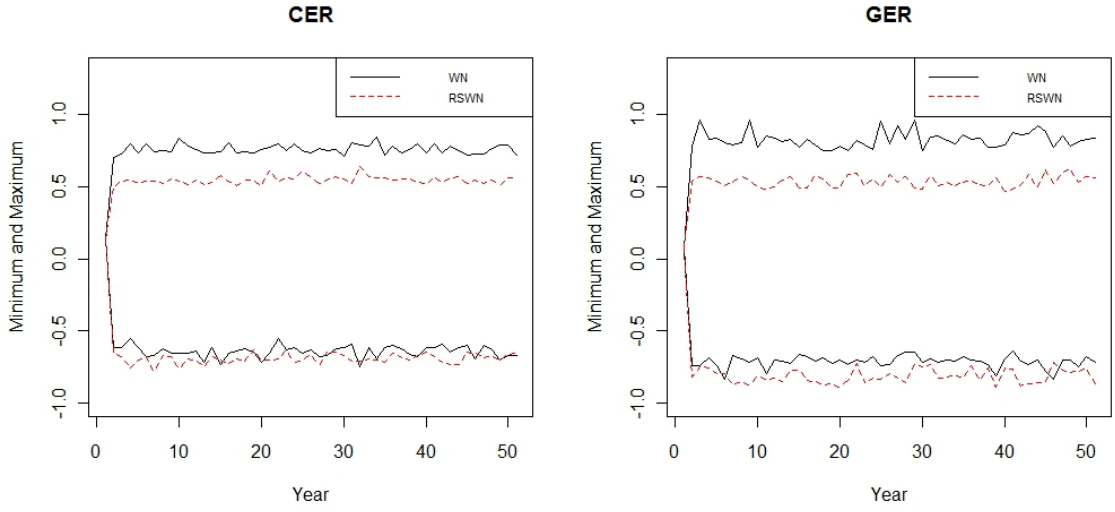

Figure 5.11: WN/RSWN Models Minimum and Maximum of Projection for the Forces of Annual Canadian Equity Return and Global Equity Return

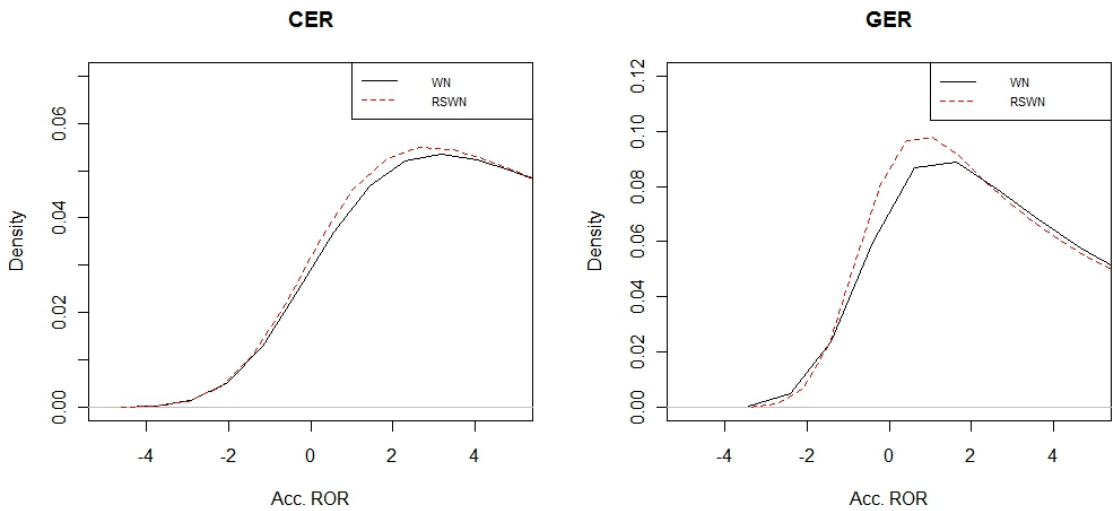

Figure 5.12: WN/RSWN Models Distribution of Accumulated Rates of Return at Year 50 of Projection for the Forces of Annual Canadian Equity Return and Global Equity Return

In conclusion, the RSWN model is preferable to the WN model for pension plan projections. Note that extreme values analysis is not performed for equity returns because large increases or decreases in stock prices are not rare. Also, there are no effective mechanisms in the market or from the Government to prevent large fluctuations in stock prices.

Chapter 6

Pension Plan Projections

6.1 Methodology

For the projections of the DB pension plan and DC pension plan, 100,000 trials are simulated using the final models selected for inflation, wage index, long-term interest rate, Canadian equity return, and global equity return in the previous chapters.

6.1.1 Defined Benefit Pension Plan

A DB pension plan projection is essentially a series of valuations at each valuation year (this ranges from one year to three years) in the future. The common practice for consulting actuaries performing pension plan projections is to use deterministic variables based on the values of the valuation variables. In other words, the same value is used for every future valuation. To investigate sensitivity of an economic variable, the value is changed deterministically to see how the results change. That is, the change is only from one single number to another single number. For example, if the discount rate used for the last valuation is 5% per year, the discount rate for the projection is also 5%. To investigate the impact of discount rate being lower, the projection is performed again using, say, a 4.5% discount rate. Then the results between two projections are compared to determine the impact. This method does not consider the stochastic nature of economic series.

In this report, the steps of the pension plan projection are the same as above with the exception that the economic variables are now stochastic. This means the economic series are allowed to change in the future; however, the simulated values for the economic

variables are used for future valuations as if they are deterministic as required by the pension legislations in Canada. For example, if the discount rate in year one is 5%, the valuation performed at the end of year one uses a 5% discount rate. In year two, if the discount rate changes to 5.5%, the discount rate for valuation performed at the end of year two is 5.5%.

For illustrations, this report:

- studies a pension plan with only one member. A typical pension plan has at least 10 members with different demographic profiles;
- assumes the member retires at age 60. A typical valuation also assumes termination of employment, death, and retirement at different ages with different probabilities;
- ignores the post-retirement period;
- uses the total unit credit (TUC) funding method instead of the projected unit credit (PUC) funding method as required by the pension legislations. For details of TUC and PUC, see Aitken (1996);
- sets the discount rate at the prevailing simulated interest rate only. In reality, the discount rate should be the expected future return on the pension assets adjusted by the additional expected return from active management, investment expenses, and a margin for adverse deviation;
- performs going-concern valuations only; pension legislations also require solvency valuations;
- assumes that the contribution to fund the accumulation of the benefit (called normal cost) continues regardless of the going-concern surplus level. However, in Canada, the *Income Tax Act* prevents the employer from contributing when the pension assets are too large compared to the pension liabilities;
- amortizes any going-concern deficit over three years, and the amount to be amortized in the next valuation is recalculated solely based on the new funded status. In reality, contributions for the deficit (called special payments) are determined based on a schedule where each deficit has its own contributions. Also, the amortization period for a going-concern deficit is typically 15 years; and

- performs a valuation every year. Depending on the jurisdiction in Canada, valuation may be required to be performed every one to three years. The frequency of valuation also depends on the funded status at the last valuation.

Mortality rates are assumed to follow the 1994 Uninsured Pensioners Mortality Table (UP94). This is a fairly common mortality table for pension plan valuations. In recent years, there has been an expectation to project the mortality rates using AA scale. Therefore, the mortality rates used for the pension plan projection follow UP94 projected to year 2020 (denoted UP94@2020). The original UP94 rates and AA scale are summarized in Appendix A. Only the rates from age 60 and onwards are shown because it is assumed that the member retires at age 60 and there are no pre-retirement deaths.

6.1.2 Defined Contribution Pension Plan

There is no legislation requiring a valuation or a projection for DC pension plans in Canada. However, most of the DC pension plan administration companies provide some sort of planning tool which involves simple deterministic projection of the account balance.

In this report, the account balance is projected forward to retirement using stochastic returns (as a weighted average of long-term interest rate, Canadian equity return, and global equity return). The account balance is then converted into a pension using the prevailing interest rate at retirement. The replacement ratio is calculated as the ratio of the pension amount to the pre-retirement wages. Replacement ratio is a better measure of benefit adequacy than just the account balance at retirement because it is a measure of standard of living.

For illustrations, this report:

- assumes the asset mix is the same for each year. However, some financial institutions offer life-cycle products which increase the weighting of fixed-income as the member grows older;

- uses the long-term interest rate to calculate the pension. However, this is not always the case for insurance companies;

- uses the UP94@2020 mortality rates. In reality, insurance companies usually have their own mortality tables which can be significantly different than the UP94 mortality table; and

- ignores any loading added to the purchase price of an annuity. In reality, the price charged by the insurance companies includes a loading for expenses and profit.

6.1.3 Demographic Variables and Plan Provisions

Table 6.1 summarizes the demographic variables and the plan provisions for the pension plan projections in this report. Note that demographic variables are assumed to be deterministic.

Table 6.1: Demographic variables and Plan Provisions of the Sample DB and DC Pension Plans

Plan	DB	DC
Hire Age	30	30
Sex	Male	Male
Starting Salary	50,000	50,000
Asset Mix - Fixed-income - Canadian Equity - Global Equity	 40% 30% 30%	 40% 30% 30%
Retirement Age	60	60
Termination of Employment	Nil	N/A
Pre-retirement Death	Nil	N/A
Post-retirement Death	UP94@2020	UP94@2020
Plan Provisions	2% of final wages times years of service	Employer contributes 10% of wages annually

6.2 Projection Results

6.2.1 Defined Benefit Pension Plan

Figures 6.1-6.6 show the mean, standard deviation, 5th and 95th percentiles, and minimum and maximum for surplus, employer contributions, and employer contributions as a percentage of wages.

6.2.2 Defined Contribution Pension Plan

Figures 6.7-6.9 show the histogram of account balance at retirement, pension purchased using the account balance, and the replacement ratio. Table 6.2 summarizes the statistics of the distribution of the three indicator items above.

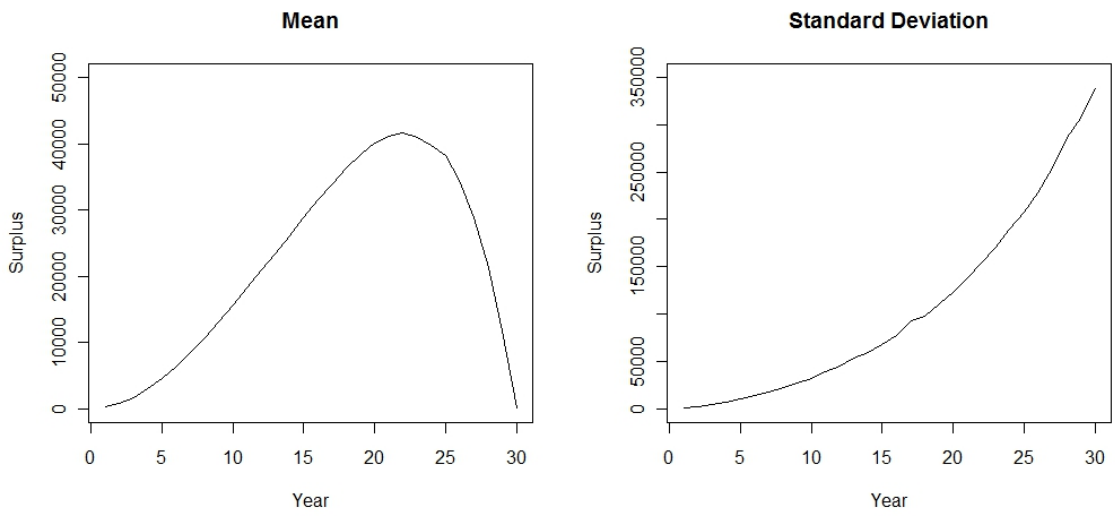

Figure 6.1: Mean and Standard Deviation of Projected DB Surplus

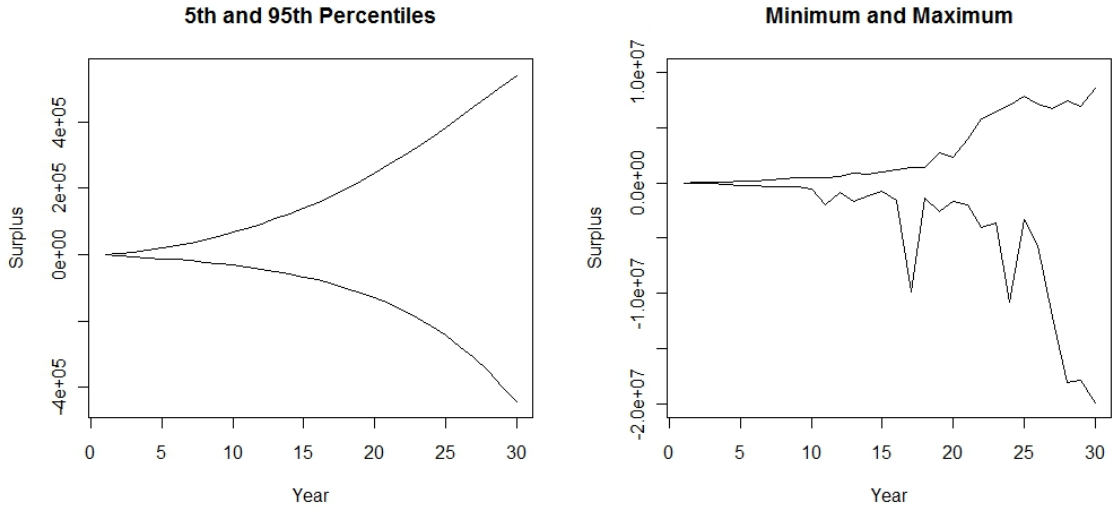

Figure 6.2: 5th and 95th Percentiles and Minimum and Maximum of Projected DB Surplus

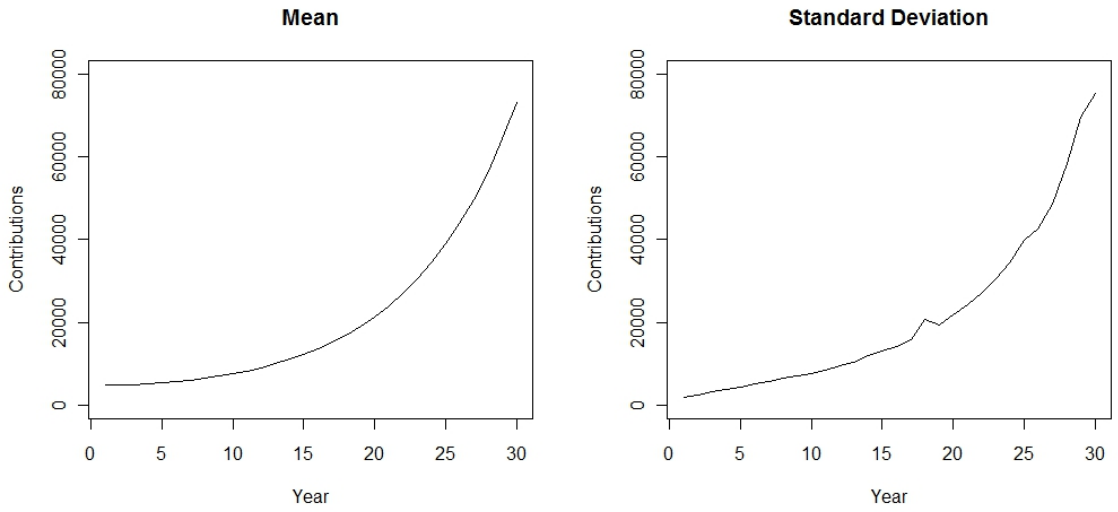

Figure 6.3: Mean and Standard Deviation of Projected DB Employer Contribution

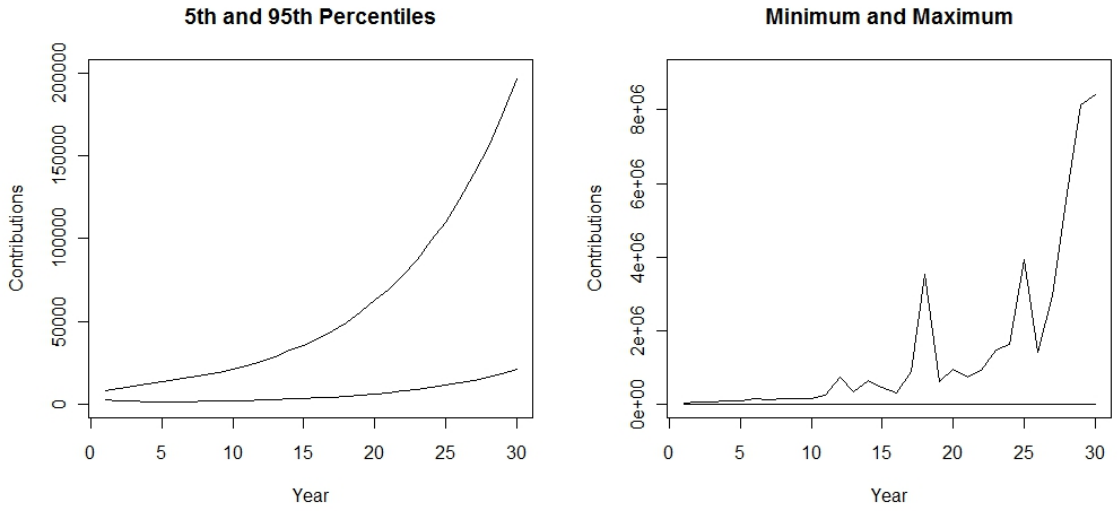

Figure 6.4: 5th and 95th Percentiles and Minimum and Maximum of Projected DB Employer Contribution

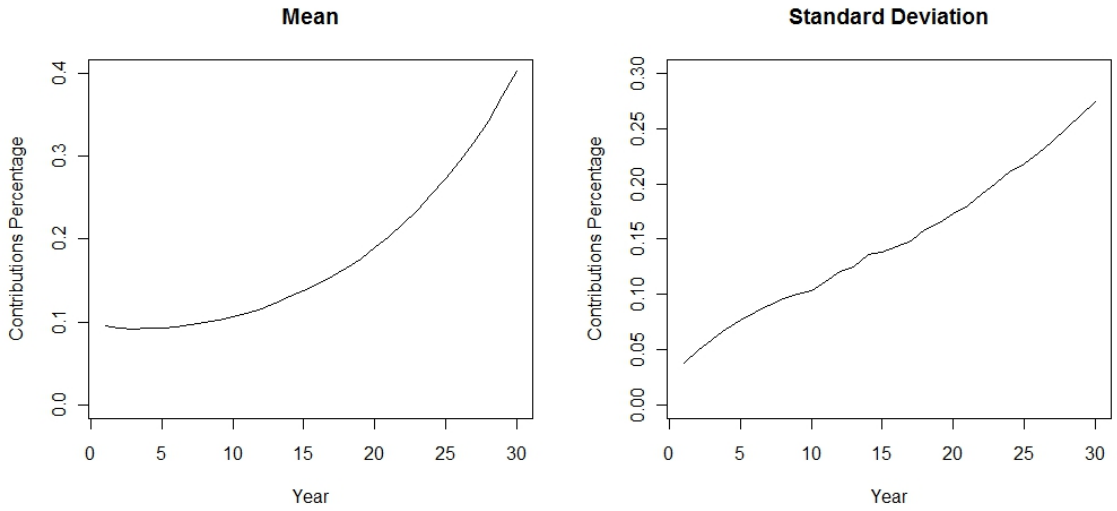

Figure 6.5: Mean and Standard Deviation of Projected DB Employer Contribution Percentage

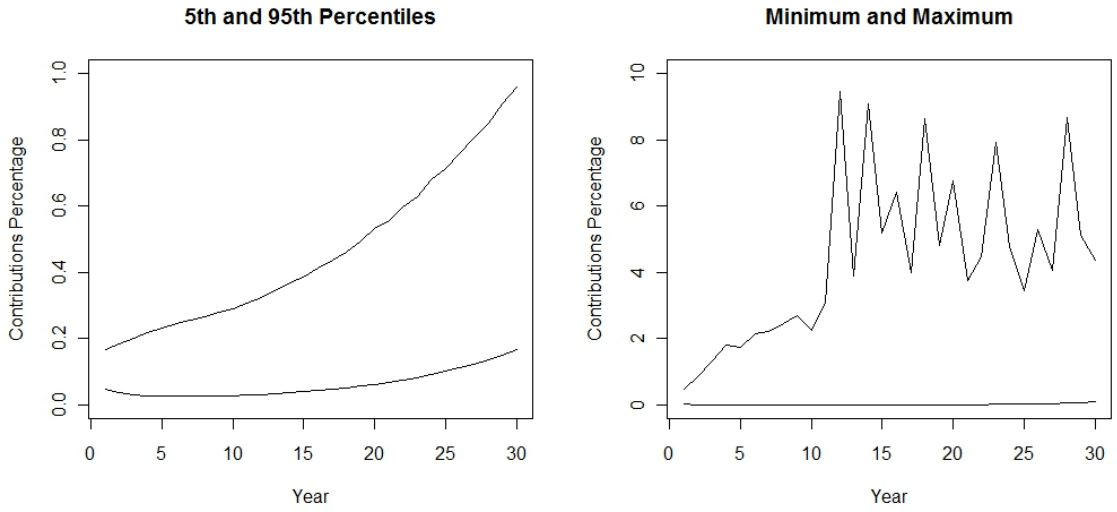

Figure 6.6: 5th and 95th Percentiles and Minimum and Maximum of Projected DB Employer Contribution Percentage

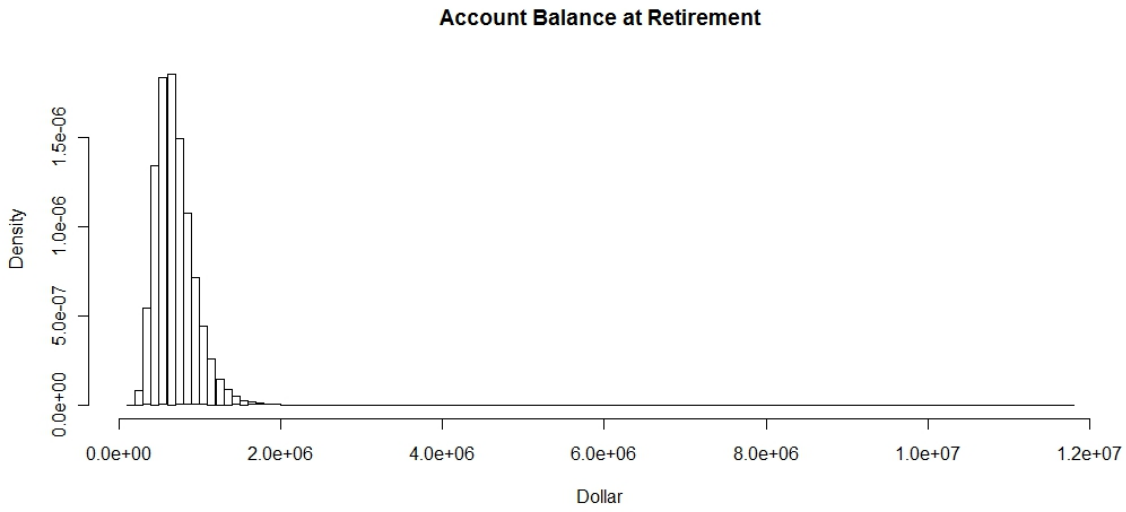

Figure 6.7: Histogram of Projected DC Account Balance at Retirement

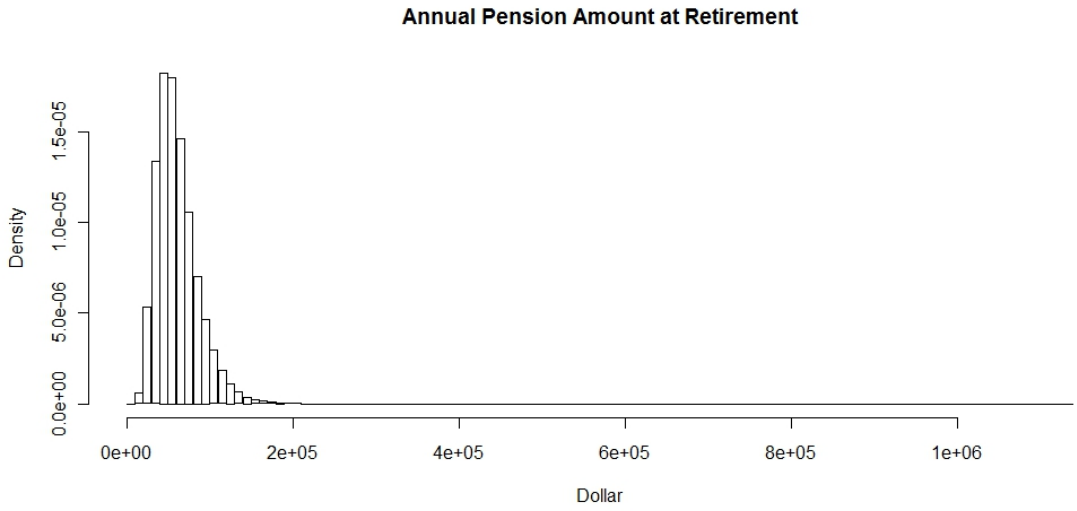

Figure 6.8: Histogram of Projected Annual Pension Amount at Retirement Purchased using DC Account Balance

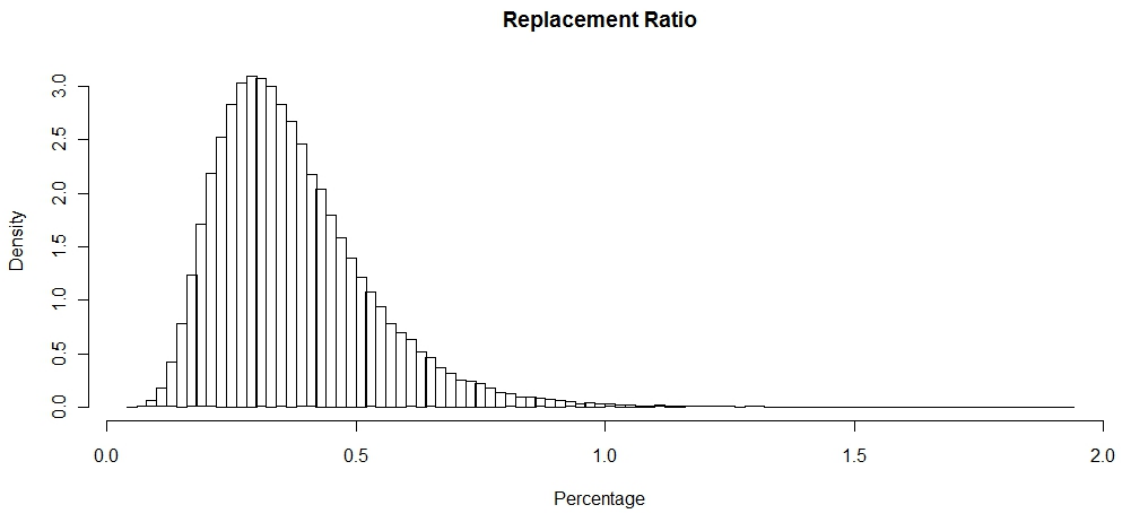

Figure 6.9: Histogram of Projected DC Replacement Ratio

6.3 Comments

6.3.1 Defined Benefit Pension Plan

The standard deviation of surplus increases over 30 years of the employee's career. This is expected because the randomness of the economic series gets compounded every year. The increase in standard deviation can also be observed in Figure 6.2 where the gap between the 5th and 95th percentiles increases over time. Notice that the percentiles are not symmetric around zero. The 95th percentile is larger in magnitude than the 5th percentile. However, minimum and maximum is the opposite where minimum values tend to be larger in magnitude than maximum values. This suggests a skewed distribution for surplus. Interestingly, the mean of surplus increases for 22 years and then decreases to almost zero at year 30.

The mean of the employer contributions goes up every year. This is expected because, apart from the randomness, normal costs go up every year as the member ages. Also, TUC does not take into account the future wage increases. Moreover, special payments can only be positive. That is, the surplus accumulated in the assets cannot be used to pay for the normal cost while the deficit needs to be amortized in the form of special payment. The non-symmetry of the employer contributions clearly shows in Figure 6.4 where the 5th percentile and minimum value are close to the normal costs and the 95th percentile and maximum

Table 6.2: Statistics of Projected Data for DC Indicator Items

Indicator Items	Account Balance	Pension Amount	Replacement Ratio
Mean	700,458	61,391	37.46%
SD	244,445	26,195	15.65%
Minimum	167,385	6,672	5.30%
5th Percentile	384,910	28,993	17.77%
25th Percentile	530,867	43,257	26.39%
Median	662,796	56,800	34.60%
75th Percentile	828,207	74,337	45.34%
95th Percentile	1,139,557	108,813	66.60%
Maximum	11,750,390	1,134,736	193.26%

values are unbounded.

In Figures 6.5 and 6.6, the graphs of the statistics of the employer contribution as a percentage of wages are similar to the ones for employer contribution. The only difference between these two items is the additional step of dividing the employer contributions by wages. This step has offset some of the randomness introduced by the wages as illustrated in Figure 6.6 where the 95th percentile of the contribution percentage did not grow as fast as the 95th percentile of the contribution. Wages determine the pension amount and in turn the pension liabilities. The special payment is determined by the surplus which is simply the difference between pension assets and liabilities. Therefore, fluctuations in wages affect the distribution of the employer contributions. However, by dividing the employer contributions by wages, part of the effect of fluctuations is canceled out in the employer contribution percentage.

Employers in DB pension plans are most interested in employer contributions as a percentage of wages. A mean contribution percentage of over 20% of wages is quite high. On top of that, 5% of the time the employer has to contribute 100% of wages, and, in the worst case scenario, close to 1000% of wages. This simple projection highlights the high volatility of the employer contributions, which is a common belief among employers.

6.3.2 Defined Contribution Pension Plan

The distribution of the account balance at retirement is very skewed as illustrated in Figure 6.7 and in Table 6.2. Notice the account balance is always positive among the 100,000 trials. This is because the employer contributes into the account every year. Also, the accumulated

amount is positively skewed as shown in Figure 5.12.

The distribution of the pension purchased by the account balance has a very similar shape as the distribution of the account balance. This is because the pension purchased is the account balance divided by a life annuity factor calculated using the long-term interest rate at retirement. The minimum pension amount is about $6,700. But this amount does not provide information regarding the standard of living without comparing it to the wages just before retirement.

The distribution of the replacement ratio is still very skewed but less skewed than the distribution of the pension purchased. Again, similar to the employer contribution percentage in Section 6.3.1, dividing the pension by the wages eliminates some volatility in the replacement ratios. Also note that the path of the wage increases affects the account balance, all else being equal, because larger increases in the earlier stage of the career means more time for the larger contributions to earn interests to retirement.

DC pension plan employees are most interested in the replacement ratio. In general, a retiree needs over 70% replacement ratio to maintain the pre-retirement living standards (Aon Consulting, 2008). Diamond (1977) suggested that a replacement ratio of 2/3 from a pension plan is adequate. Therefore, out of the 70%, more than 60% should come from employer-sponsored pension plan, government pension plan, and social security. As government pension plan and social security are generally not substantial, the majority of the 60% should come from employer-sponsored pension plan. The mean replacement ratio is 37.46% which is significantly below 60%. The employees have to be in almost the 95th percentile to get the 60% replacement ratio. Moreover, the worst case scenario is as low as 5.30% of pre-retirement wages. Obviously, this poses a problem to maintain not only the pre-retirement living standards but also the minimum living standards. In addition, the impact of inflation has not been taken into account. The pension amount may be adequate at retirement; however, the purchasing power of the pension amount may not be adequate 10 years into retirement if the inflation has been high since retirement.

This simple DC pension plan projection highlights the inadequacy of retirement income for employees participating in DC pension plans. As employees have not yet retired in a large group (like the baby boomers) from DC pension plans, the fact that DC pension plan does not provide enough income for retirement has not been felt by the society yet. In 10 years times, government and citizens may rethink the vehicles to provide retirement income that balance the shortcomings of both DB pension plans and DC pension plans.

Chapter 7

Conclusions

This project focused on key economic variables for pension plan projections. A class of time series models was selected to model the economic variables. After parameter estimation, simulations were performed with different initial values. Then the final model for each variable is selected using multiple criteria: parameter estimates, statistics of the simulated data for the next 50 years, dynamics of the model, extreme values, economic theories, and common actuarial practices. The final models are:

- inflation: regime switching ARCH(1) proportional model;
- wage index: transfer function of lag two with inflation as an exogenous variable;
- long-term interest rate: transfer function of lag two with inflation as an exogenous variable; and
- Canadian equity return and global equity return: regime switching vector white noise model.

The selected final models and parameter estimates are used to simulate a sample DB pension plan and a sample DC pension plan, and brief comments were made for each projection. The projection highlights the high volatility of employer contributions in DB pension plans and the inadequacy of retirement income in DC pension plans.

In future research, several areas can be investigated. For example, consider

- the criteria for the parameters so that the simulated data is always positive for the ARCH(P) model similar to the one for Cox-Ingersoll-Ross model (Cox, Ingersoll &

Ross, 1985);

- statistical tests to rank the different time series models including regime switching models and transfer function models;
- the use of Bayesian model averaging instead of selecting a single model for each variable;
- the use of stochastic demographic variables;
- a more realistic approach to DB pension plan projection;
- an expansion of the asset classes in the asset mix (as well as more time series models to model the new asset classes);
- an expansion of the types of pension plans to include hybrid pension plans (i.e. pension plans with both DB and DC components);
- accounting valuations including the calculation of pension expenses;
- the impact of changes in asset mix;
- the impact of changes in benefit provisions of a DB pension plan; and
- the impact of changes in the employer contribution rate of a DC pension plan.

Appendix A

Mortality Table

The males mortality rates and AA projection scales for the 1994 Uninsured Pensioner Mortality Table are summarized in the following table.

Source: Society of Actuaries UP-94 Task Force. (1995). The 1994 uninsured pensioner mortality table. *Transactions of the Society of Actuaries.* 47.

Age	Rate	AA Scale	Age	Rate	AA Scale
60	0.008576	0.016	90	0.164442	0.004
61	0.009663	0.015	91	0.179849	0.004
62	0.010911	0.015	92	0.196001	0.003
63	0.012335	0.014	93	0.213325	0.003
64	0.013914	0.014	94	0.231936	0.003
65	0.015629	0.014	95	0.251189	0.002
66	0.017462	0.013	96	0.270441	0.002
67	0.019391	0.013	97	0.289048	0.002
68	0.021354	0.014	98	0.306750	0.001
69	0.023364	0.014	99	0.323976	0.001
70	0.025516	0.015	100	0.341116	0.001
71	0.027905	0.015	101	0.358560	0.000
72	0.030625	0.015	102	0.376699	0.000
73	0.033549	0.015	103	0.396884	0.000
74	0.036614	0.015	104	0.418855	0.000
75	0.040012	0.014	105	0.440585	0.000
76	0.043933	0.014	106	0.460043	0.000
77	0.048570	0.013	107	0.475200	0.000
78	0.053991	0.012	108	0.485670	0.000
79	0.060066	0.011	109	0.492807	0.000
80	0.066696	0.010	110	0.497189	0.000
81	0.073780	0.009	111	0.499394	0.000
82	0.081217	0.008	112	0.500000	0.000
83	0.088721	0.008	113	0.500000	0.000
84	0.096358	0.007	114	0.500000	0.000
85	0.104559	0.007	115	0.500000	0.000
86	0.113755	0.007	116	0.500000	0.000
87	0.124377	0.006	117	0.500000	0.000
88	0.136537	0.005	118	0.500000	0.000
89	0.149949	0.005	119	0.500000	0.000
			120	1.000000	0.000

Appendix B

Models: Definitions, Log-likelihood Functions, and Properties

This appendix provides the formulas for the definition, log-likelihood function, and the first and second derivatives of the log-likelihood function of the models if they are determinable. The formulas for the stationary mean and stationary variance are also produced. Note that, for regime switching models, only two regimes are considered. The log-likelihood functions for regime switching models are obtained using the filtering method proposed by Hamilton (1989).

B.1 White Noise Model

Definition

$X(t) = \mu + \epsilon(t) \qquad$ where $\epsilon(t)$ i.i.d. $N\left(0, \sigma^2\right)$

B.1.1 Log-likelihood Function

$$l(\boldsymbol{\theta}) = -nln\sigma - \frac{n}{2}ln(2\pi) - \frac{1}{2\sigma^2}\sum_{t=0}^{n-1}[x(t) - \mu]^2$$

B.1.2 First Derivative of the Log-likelihood Function

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} = \frac{1}{\sigma^2}\sum_{t=0}^{n-1}[x(t)-\mu]$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{t=0}^{n-1}[x(t)-\mu]^2$$

B.1.3 Second Derivative of the Log-likelihood Function

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu^2} = -\frac{n}{\sigma^2}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \sigma} = -\frac{2}{\sigma^3}\sum_{t=0}^{n-1}[x(t)-\mu]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3}{\sigma^4}\sum_{t=0}^{n-1}[x(t)-\mu]^2$$

B.1.4 Stationary Mean and Stationary Variance

$$E[X(t)] = \mu$$

$$Var(X(t)) = \sigma^2$$

B.2 AR(1) Model

Definition

$$X(t) = \mu + \alpha[X(t-1)-\mu] + \epsilon(t) \qquad \text{where } \epsilon(t) \text{ i.i.d. } N\left(0, \sigma^2\right)$$

B.2.1 Log-likelihood Function

By conditioning on $X(0)$, one gets

$$l(\boldsymbol{\theta}) = -(n-1)ln\sigma - \frac{n-1}{2}ln(2\pi) - \frac{1}{2\sigma^2}\sum_{t=1}^{n-1}[x(t)-\mu-\alpha x(t-1)+\alpha\mu]^2$$

B.2.2 First Derivative of the Log-likelihood Function

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} = \frac{1-\alpha}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - \mu - \alpha x(t-1) + \alpha\mu]$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} = \frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - \mu - \alpha x(t-1) + \alpha\mu][x(t-1) - \mu]$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \sigma} = -\frac{n-1}{\sigma} + \frac{1}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - \mu - \alpha x(t-1) + \alpha\mu]^2$$

B.2.3 Second Derivative of the Log-likelihood Function

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu^2} = -\frac{(n-1)(1-\alpha)^2}{\sigma^2}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \alpha} = -\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - \mu - \alpha x(t-1) + \alpha\mu] - \frac{1-\alpha}{\sigma^2}\sum_{t=1}^{n-1}[x(t-1) - \mu]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \sigma} = -\frac{2(1-\alpha)}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - \mu - \alpha x(t-1) + \alpha\mu]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha^2} = -\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t-1) - \mu]^2$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \sigma} = -\frac{2}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - \mu - \alpha x(t-1) + \alpha\mu][x(t-1) - \mu]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma^2} = \frac{n-1}{\sigma^2} - \frac{3}{\sigma^4}\sum_{t=1}^{n-1}[x(t) - \mu - \alpha x(t-1) + \alpha\mu]^2$$

B.2.4 Stationary Mean and Stationary Variance

$$E[X(t)] = \mu$$

$$Var(X(t)) = \frac{\sigma^2}{1-\alpha^2} \quad \text{if} \quad \alpha^2 < 1$$

B.3 ARCH(1) Model - Full

Definition

$X(t) = \mu + \alpha\left[X(t-1) - \mu\right] + \epsilon(t)$ where $\epsilon(t) \sim N\left(0, \sigma^2(t)\right)$ and
$\sigma^2(t) = \lambda^2 + \gamma^2\left[X(t-1) - \mu\right]^2$

B.3.1 Log-likelihood Function

By conditioning on $X(0)$, one gets

$$\begin{aligned} l(\boldsymbol{\theta}) = &-\frac{1}{2}\sum_{t=1}^{n-1} ln\left(\lambda^2 + \gamma^2\left[x(t-1) - \mu\right]^2\right) - \frac{n-1}{2} ln(2\pi) \\ &- \frac{1}{2}\sum_{t=1}^{n-1} \frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]^2}{\lambda^2 + \gamma^2\left[x(t-1) - \mu\right]^2} \end{aligned}$$

B.3.2 First Derivative of the Log-likelihood Function

$$\begin{aligned} \frac{\partial l(\boldsymbol{\theta})}{\partial \mu} = &\gamma^2 \sum_{t=1}^{n-1} \frac{x(t-1) - \mu}{\lambda^2 + \gamma^2\left[x(t-1) - \mu\right]^2} \\ &+ (1-\alpha)\sum_{t=1}^{n-1} \frac{x(t) - \mu - \alpha x(t-1) + \alpha\mu}{\lambda^2 + \gamma^2\left[x(t-1) - \mu\right]^2} \\ &- \gamma^2 \sum_{t=1}^{n-1} \frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]^2\left[x(t-1) - \mu\right]}{\left(\lambda^2 + \gamma^2\left[x(t-1) - \mu\right]^2\right)^2} \end{aligned}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \lambda} = -\lambda \sum_{t=1}^{n-1} \frac{1}{\lambda^2 + \gamma^2\left[x(t-1) - \mu\right]^2} + \lambda \sum_{t=1}^{n-1} \frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]^2}{\left(\lambda^2 + \gamma^2\left[x(t-1) - \mu\right]^2\right)^2}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} = \sum_{t=1}^{n-1} \frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]\left[x(t-1) - \mu\right]}{\lambda^2 + \gamma^2\left[x(t-1) - \mu\right]^2}$$

$$\begin{aligned} \frac{\partial l(\boldsymbol{\theta})}{\partial \gamma} = &-\gamma \sum_{t=1}^{n-1} \frac{\left[x(t-1) - \mu\right]^2}{\lambda^2 + \gamma^2\left[x(t-1) - \mu\right]^2} \\ &+ \gamma \sum_{t=1}^{n-1} \frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]^2\left[x(t-1) - \mu\right]^2}{\left(\lambda^2 + \gamma^2\left[x(t-1) - \mu\right]^2\right)^2} \end{aligned}$$

B.3.3 Second Derivative of the Log-likelihood Function

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu^2} = & -\gamma^2 \sum_{t=1}^{n-1} \frac{1}{\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2} + 2\gamma^4 \sum_{t=1}^{n-1} \frac{\left[x(t-1) - \mu\right]^2}{\left(\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2\right)^2} \\
& - (1-\alpha)^2 \sum_{t=1}^{n-1} \frac{1}{\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2} \\
& + 4(1-\alpha)\gamma^2 \sum_{t=1}^{n-1} \frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]\left[x(t-1) - \mu\right]}{\left(\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2\right)^2} \\
& + \gamma^2 \sum_{t=1}^{n-1} \frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]^2}{\left(\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2\right)^2} \\
& - 4\gamma^4 \sum_{t=1}^{n-1} \frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]^2 \left[x(t-1) - \mu\right]^2}{\left(\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2\right)^3}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \lambda} = & -2\lambda\gamma^2 \sum_{t=1}^{n-1} \frac{x(t-1) - \mu}{\left(\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2\right)^2} \\
& - 2\lambda(1-\alpha) \sum_{t=1}^{n-1} \frac{x(t) - \mu - \alpha x(t-1) + \alpha\mu}{\left(\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2\right)^2} \\
& + 4\lambda\gamma^2 \sum_{t=1}^{n-1} \frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]^2 \left[x(t-1) - \mu\right]}{\left(\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2\right)^3}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \alpha} = & -\sum_{t=1}^{n-1} \frac{x(t) - \mu - \alpha x(t-1) + \alpha\mu}{\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2} \\
& - (1-\alpha) \sum_{t=1}^{n-1} \frac{x(t-1) - \mu}{\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2} \\
& + 2\gamma^2 \sum_{t=1}^{n-1} \frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]\left[x(t-1) - \mu\right]^2}{\left(\lambda^2 + \gamma^2 \left[x(t-1) - \mu\right]^2\right)^2}
\end{aligned}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\mu\partial\gamma} = 2\gamma\sum_{t=1}^{n-1}\frac{x(t-1)-\mu}{\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2} - 2\gamma^3\sum_{t=1}^{n-1}\frac{\left[x(t-1)-\mu\right]^3}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^2}$$
$$- 2\left(1-\alpha\right)\gamma\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]\left[x(t-1)-\mu\right]^2}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^2}$$
$$- 2\gamma\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]^2\left[x(t-1)-\mu\right]}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^2}$$
$$+ 4\gamma^3\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]^2\left[x(t-1)-\mu\right]^3}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^3}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\lambda^2} = -\sum_{t=1}^{n-1}\frac{1}{\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2} + 2\lambda^2\sum_{t=1}^{n-1}\frac{1}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^2}$$
$$+ \sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]^2}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^2}$$
$$- 4\lambda^2\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]^2}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^3}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\lambda\partial\alpha} = -2\lambda\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]\left[x(t-1)-\mu\right]}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^2}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\lambda\partial\gamma} = 2\lambda\gamma\sum_{t=1}^{n-1}\frac{\left[x(t-1)-\mu\right]^2}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^2}$$
$$- 4\lambda\gamma\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]^2\left[x(t-1)-\mu\right]^2}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^3}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha^2} = -\sum_{t=1}^{n-1}\frac{\left[x(t-1)-\mu\right]^2}{\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha\partial\gamma} = -2\gamma\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]\left[x(t-1)-\mu\right]^3}{\left(\lambda^2+\gamma^2\left[x(t-1)-\mu\right]^2\right)^2}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma^2} = -\sum_{t=1}^{n-1} \frac{[x(t-1)-\mu]^2}{\lambda^2+\gamma^2[x(t-1)-\mu]^2} + 2\gamma^2 \sum_{t=1}^{n-1} \frac{[x(t-1)-\mu]^4}{\left(\lambda^2+\gamma^2[x(t-1)-\mu]^2\right)^2}$$
$$+ \sum_{t=1}^{n-1} \frac{[x(t)-\mu-\alpha x(t-1)+\alpha\mu]^2[x(t-1)-\mu]^2}{\left(\lambda^2+\gamma^2[x(t-1)-\mu]^2\right)^2}$$
$$- 4\gamma^2 \sum_{t=1}^{n-1} \frac{[x(t)-\mu-\alpha x(t-1)+\alpha\mu]^2[x(t-1)-\mu]^4}{\left(\lambda^2+\gamma^2[x(t-1)-\mu]^2\right)^3}$$

B.3.4 Stationary Mean and Stationary Variance

$E[X(t)] = \mu$

$$\begin{aligned} Var(X(t)) &= E[Var(X(t)|X(t-1))] + Var(E[X(t)|X(t-1)]) \\ &= E[\sigma^2(t)] + Var(\mu + \alpha[X(t-1)-\mu]) \\ &= \lambda^2 + \gamma^2 E[[X(t-1)-\mu]^2] + \alpha^2 Var(X(t-1)) \\ &= \lambda^2 + \gamma^2 Var(X(t)) + \alpha^2 Var(X(t)) \end{aligned}$$

$$\Rightarrow \qquad Var(X(t)) = \frac{\lambda^2}{1-\alpha^2-\gamma^2} \quad \text{if} \quad \alpha^2+\gamma^2 < 1$$

B.4 ARCH(1) Model - Non-Centered

Definition

$X(t) = \mu + \alpha[X(t-1)-\mu] + \epsilon(t)$ where $\epsilon(t) \sim N\left(0, \sigma^2(t)\right)$ and
$\sigma^2(t) = \lambda^2 + \gamma^2 X^2(t-1)$

B.4.1 Log-likelihood Function

By conditioning on $X(0)$, one gets

$$l(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{t=1}^{n-1} ln\left[\lambda^2+\gamma^2 x^2(t-1)\right] - \frac{n-1}{2} ln(2\pi)$$
$$- \frac{1}{2}\sum_{t=1}^{n-1} \frac{[x(t)-\mu-\alpha x(t-1)+\alpha\mu]^2}{\lambda^2+\gamma^2 x^2(t-1)}$$

B.4.2 First Derivative of the Log-likelihood Function

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} = (1-\alpha)\sum_{t=1}^{n-1}\frac{x(t)-\mu-\alpha x(t-1)+\alpha\mu}{\lambda^2+\gamma^2x^2(t-1)}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \lambda} = -\lambda\sum_{t=1}^{n-1}\frac{1}{\lambda^2+\gamma^2x^2(t-1)} + \lambda\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]^2}{\left[\lambda^2+\gamma^2x^2(t-1)\right]^2}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} = \sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]\left[x(t-1)-\mu\right]}{\lambda^2+\gamma^2x^2(t-1)}$$

$$\begin{aligned}\frac{\partial l(\boldsymbol{\theta})}{\partial \gamma} &= -\gamma\sum_{t=1}^{n-1}\frac{x^2(t-1)}{\lambda^2+\gamma^2x^2(t-1)} \\ &\quad + \gamma\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]^2x^2(t-1)}{\left[\lambda^2+\gamma^2x^2(t-1)\right]^2}\end{aligned}$$

B.4.3 Second Derivative of the Log-likelihood Function

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu^2} = -(1-\alpha)^2\sum_{t=1}^{n-1}\frac{1}{\lambda^2+\gamma^2x^2(t-1)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu\partial \lambda} = -\lambda(1-\alpha)\sum_{t=1}^{n-1}\frac{x(t)-\mu-\alpha x(t-1)+\alpha\mu}{\left[\lambda^2+\gamma^2x^2(t-1)\right]^2}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu\partial \alpha} = -\sum_{t=1}^{n-1}\frac{x(t)-\mu-\alpha x(t-1)+\alpha\mu}{\lambda^2+\gamma^2x^2(t-1)} - (1-\alpha)\sum_{t=1}^{n-1}\frac{x(t-1)-\mu}{\lambda^2+\gamma^2x^2(t-1)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu\partial \gamma} = -2(1-\alpha)\gamma\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]x^2(t-1)}{\left[\lambda^2+\gamma^2x^2(t-1)\right]^2}$$

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda^2} &= -\sum_{t=1}^{n-1}\frac{1}{\left[\lambda^2+\gamma^2x^2(t-1)\right]^2} + 2\lambda^2\sum_{t=1}^{n-1}\frac{1}{\left[\lambda^2+\gamma^2x^2(t-1)\right]^2} \\ &\quad + \sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]^2}{\left[\lambda^2+\gamma^2x^2(t-1)\right]^2} \\ &\quad - 4\lambda^2\sum_{t=1}^{n-1}\frac{\left[x(t)-\mu-\alpha x(t-1)+\alpha\mu\right]^2}{\left[\lambda^2+\gamma^2x^2(t-1)\right]^3}\end{aligned}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\lambda\partial\alpha} = -2\lambda\sum_{t=1}^{n-1}\frac{[x(t)-\mu-\alpha x(t-1)+\alpha\mu]\,[x(t-1)-\mu]}{[\lambda^2+\gamma^2x^2(t-1)]^2}$$

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\lambda\partial\gamma} &= 2\lambda\gamma\sum_{t=1}^{n-1}\frac{x^2(t-1)}{[\lambda^2+\gamma^2x^2(t-1)]^2}\\ &\quad - 4\lambda\gamma\sum_{t=1}^{n-1}\frac{[x(t)-\mu-\alpha x(t-1)+\alpha\mu]^2\,x^2(t-1)}{[\lambda^2+\gamma^2x^2(t-1)]^3}\end{aligned}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha^2} = -\sum_{t=1}^{n-1}\frac{[x(t-1)-\mu]^2}{\lambda^2+\gamma^2x^2(t-1)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha\partial\gamma} = -2\gamma\sum_{t=1}^{n-1}\frac{[x(t)-\mu-\alpha x(t-1)+\alpha\mu]\,[x(t-1)-\mu]\,x^2(t-1)}{[\lambda^2+\gamma^2x^2(t-1)]^2}$$

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\gamma^2} &= -\sum_{t=1}^{n-1}\frac{x^2(t-1)}{\lambda^2+\gamma^2x^2(t-1)} + 2\gamma^2\sum_{t=1}^{n-1}\frac{x^4(t-1)}{[\lambda^2+\gamma^2x^2(t-1)]^2}\\ &\quad + \sum_{t=1}^{n-1}\frac{[x(t)-\mu-\alpha x(t-1)+\alpha\mu]^2\,x^2(t-1)}{[\lambda^2+\gamma^2x^2(t-1)]^2}\\ &\quad - 4\gamma^2\sum_{t=1}^{n-1}\frac{[x(t)-\mu-\alpha x(t-1)+\alpha\mu]^2\,x^4(t-1)}{[\lambda^2+\gamma^2x^2(t-1)]^3}\end{aligned}$$

B.4.4 Stationary Mean and Stationary Variance

$E[X(t)] = \mu$

$$\begin{aligned}Var(X(t)) &= E[Var(X(t)|X(t-1))] + Var(E[X(t)|X(t-1)])\\ &= E[\sigma^2(t)] + Var(\mu+\alpha[X(t-1)-\mu])\\ &= \lambda^2+\gamma^2E[X^2(t-1)] + \alpha^2Var(X(t-1))\\ &= \lambda^2+\gamma^2Var(X(t)) + \gamma^2\mu^2 + \alpha^2Var(X(t))\end{aligned}$$

$$\Rightarrow \qquad Var(X(t)) = \frac{\lambda^2+\gamma^2\mu^2}{1-\alpha^2-\gamma^2} \quad \text{if} \quad \alpha^2+\gamma^2 < 1$$

B.5 ARCH(1) Model - Proportional

Definition

$X(t) = \mu + \alpha\left[X(t-1) - \mu\right] + \epsilon(t)$ where $\epsilon(t) \sim N\left(0, \sigma^2(t)\right)$ and $\sigma^2(t) = \gamma^2 X^2(t-1)$

B.5.1 Log-likelihood Function

By conditioning on $X(0)$, one gets

$$l(\boldsymbol{\theta}) = -(n-1)ln\gamma - \frac{n-1}{2}ln(2\pi) - \sum_{t=1}^{n-1} lnx(t-1) - \frac{1}{2\gamma^2}\sum_{t=1}^{n-1}\frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]^2}{x^2(t-1)}$$

B.5.2 First Derivative of the Log-likelihood Function

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} = \frac{1-\alpha}{\gamma^2}\sum_{t=1}^{n-1}\frac{x(t) - \mu - \alpha x(t-1) + \alpha\mu}{x^2(t-1)}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} = \frac{1}{\gamma^2}\sum_{t=1}^{n-1}\frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]\left[x(t-1) - \mu\right]}{x^2(t-1)}$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \gamma} = \frac{1}{\gamma^3}\sum_{t=1}^{n-1}\frac{\left[x(t) - \mu - \alpha x(t-1) + \alpha\mu\right]^2}{x^2(t-1)} - \frac{n-1}{\gamma}$$

B.5.3 Second Derivative of the Log-likelihood Function

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu^2} = -\frac{(1-\alpha)^2}{\gamma^2}\sum_{t=1}^{n-1}\frac{1}{x^2(t-1)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \alpha} = -\frac{1}{\gamma^2}\sum_{t=1}^{n-1}\frac{x(t) - \mu - \alpha x(t-1) + \alpha\mu}{x^2(t-1)} - \frac{1-\alpha}{\gamma^2}\sum_{t=1}^{n-1}\frac{x(t-1) - \mu}{x^2(t-1)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\mu\partial\gamma} = -\frac{2\,(1-\alpha)}{\gamma^3}\sum_{t=1}^{n-1}\frac{x(t)-\mu-\alpha x(t-1)+\alpha\mu}{x^2(t-1)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha^2} = -\frac{1}{\gamma^2}\sum_{t=1}^{n-1}\frac{[x(t-1)-\mu]^2}{x^2(t-1)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha\partial\gamma} = -\frac{2}{\gamma^3}\sum_{t=1}^{n-1}\frac{[x(t)-\mu-\alpha x(t-1)+\alpha\mu]\,[x(t-1)-\mu]}{x^2(t-1)}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\gamma^2} = -\frac{3}{\gamma^4}\sum_{t=1}^{n-1}\frac{[x(t)-\mu-\alpha x(t-1)+\alpha\mu]^2}{x^2(t-1)} + \frac{n-1}{\gamma^2}$$

B.5.4 Stationary Mean and Stationary Variance

$$E[X(t)] = \mu$$

$$Var(X(t)) = \frac{\gamma^2\mu^2}{1-\alpha^2-\gamma^2} \quad \text{if} \quad \alpha^2+\gamma^2 < 1$$

B.6 Regime Switching Vector White Noise Model

Definition

Let S_t denote the regime at time t with transition matrix $P = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}$.

$\boldsymbol{X}(t) = \boldsymbol{\mu}_{S_t} + \boldsymbol{\epsilon}(t)$ where $\boldsymbol{\epsilon}(t)$ i.i.d. $N\left(\mathbf{0}, \boldsymbol{\Sigma}_{S_t}\right)$

where

$$\boldsymbol{X}(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} \qquad \boldsymbol{\mu}_{S_t} = \begin{bmatrix} \mu_{1S_t} \\ \mu_{2S_t} \end{bmatrix} \qquad \boldsymbol{\Sigma}_{S_t} = \begin{bmatrix} \sigma^2_{1S_t} & 0 \\ 0 & \sigma^2_{2S_t} \end{bmatrix}$$

B.6.1 Stationary Mean and Stationary Variance

$$\begin{cases} E\left[X_1(t)\right] = \pi_1\mu_{11} + \pi_2\mu_{12} \\ E\left[X_2(t)\right] = \pi_1\mu_{21} + \pi_2\mu_{22} \end{cases}$$

where

$$\pi_1 = \frac{\rho_{21}}{\rho_{12} + \rho_{21}} \quad \text{and} \quad \pi_2 = \frac{\rho_{12}}{\rho_{12} + \rho_{21}}$$

$$\begin{cases} Var\left(X_1(t)\right) = \pi_1\sigma_{11}^2 + \pi_2\sigma_{12}^2 + \left(\mu_{11} - \mu_{12}\right)^2 \pi_1\pi_2 \\ Var\left(X_2(t)\right) = \pi_1\sigma_{21}^2 + \pi_2\sigma_{22}^2 + \left(\mu_{21} - \mu_{22}\right)^2 \pi_1\pi_2 \end{cases}$$

B.7 Regime Switching AR(1) Model

Definition

Let S_t denote the regime at time t with transition matrix $P = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}$.

$$X(t) = \mu_{S_t} + \alpha_{S_t}[X(t-1) - \mu_{S_{t-1}}] + \epsilon(t) \qquad \text{where } \epsilon(t) \text{ i.i.d. } N\left(0, \sigma_{S_t}^2\right)$$

B.7.1 Stationary Mean and Stationary Variance

$$E[X(t)] = \pi_1\mu_1 + \pi_2\mu_2$$

$$Var(X(t)) = \pi_1 V_1 + \pi_2 V_2 + \left(\mu_1 - \mu_2\right)^2 \pi_1\pi_2$$

where V_i is the stationary variance for regime i.

At stationarity, the variance at time t for regime 1 is the variance at time t-1 for regime 1 times α_1^2 if the regime changes from 1 to 1 plus the variance at time t-1 for regime 2 times α_1^2 if the regime changes from 2 to 1 plus the variance of the innovations for regime 1. The variance at time t for regime 2 follows a similar pattern. In other words,

$$\begin{cases} V_1 = \frac{\pi_1\rho_{11}}{\pi_1}\left(\alpha_1^2 V_1 + \sigma_1^2\right) + \frac{\pi_2\rho_{21}}{\pi_1}\left(\alpha_1^2 V_2 + \sigma_1^2\right) \\ V_2 = \frac{\pi_1\rho_{12}}{\pi_2}\left(\alpha_2^2 V_1 + \sigma_2^2\right) + \frac{\pi_2\rho_{22}}{\pi_2}\left(\alpha_2^2 V_2 + \sigma_2^2\right) \end{cases}$$

Solving the system of equations, one gets

$$V_1 = \frac{\left(1-\rho_{22}\alpha_2^2\right)\sigma_1^2 + \rho_{12}\alpha_1^2\sigma_2^2}{\left(1-\rho_{11}\alpha_1^2\right)\left(1-\rho_{22}\alpha_2^2\right) - \rho_{12}\rho_{21}\alpha_1^2\alpha_2^2}$$

$$V_2 = \frac{\rho_{21}\alpha_2^2\sigma_1^2 + \left(1-\rho_{11}\alpha_1^2\right)\sigma_2^2}{\left(1-\rho_{11}\alpha_1^2\right)\left(1-\rho_{22}\alpha_2^2\right) - \rho_{12}\rho_{21}\alpha_1^2\alpha_2^2}$$

B.8 Regime Switching ARCH(1) Model - Full

Definition

Let S_t denote the regime at time t with transition matrix $P = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}$.

$X(t) = \mu_{S_t} + \alpha_{S_t}\left[X(t-1) - \mu_{S_{t-1}}\right] + \epsilon(t)$ where $\epsilon(t) \sim N\left(0, \sigma_{S_t}^2(t)\right)$ and $\sigma_{S_t}^2(t) = \lambda_{S_t}^2 + \gamma_{S_t}^2\left[X(t-1) - \mu_{S_{t-1}}\right]^2$

B.8.1 Stationary Mean and Stationary Variance

For details of derivation, see Section B.7.1.

$$E[X(t)] = \pi_1\mu_1 + \pi_2\mu_2$$

$$Var(X(t)) = \pi_1 V_1 + \pi_2 V_2 + (\mu_1 - \mu_2)^2 \pi_1\pi_2$$

where

$$\begin{cases} V_1 = \frac{\pi_1\rho_{11}}{\pi_1}\left[\left(\alpha_1^2+\gamma_1^2\right)V_1 + \lambda_1^2\right] + \frac{\pi_2\rho_{21}}{\pi_1}\left[\left(\alpha_1^2+\gamma_1^2\right)V_2 + \lambda_1^2\right] \\ V_2 = \frac{\pi_1\rho_{12}}{\pi_2}\left[\left(\alpha_2^2+\gamma_2^2\right)V_1 + \lambda_2^2\right] + \frac{\pi_2\rho_{22}}{\pi_2}\left[\left(\alpha_2^2+\gamma_2^2\right)V_2 + \lambda_2^2\right] \end{cases}$$

Solving the system of equations, one gets

$$V_1 = \frac{\left[1-\rho_{22}\left(\alpha_2^2+\gamma_2^2\right)\right]\lambda_1^2 + \rho_{12}\left(\alpha_1^2+\gamma_1^2\right)\lambda_2^2}{\left[1-\rho_{11}\left(\alpha_1^2+\gamma_1^2\right)\right]\left[1-\rho_{22}\left(\alpha_2^2+\gamma_2^2\right)\right] - \rho_{12}\rho_{21}\left(\alpha_1^2+\gamma_1^2\right)\left(\alpha_2^2+\gamma_2^2\right)}$$

$$V_2 = \frac{\rho_{21}\left(\alpha_2^2+\gamma_2^2\right)\lambda_1^2 + \left[1-\rho_{11}\left(\alpha_1^2+\gamma_1^2\right)\right]\lambda_2^2}{\left[1-\rho_{11}\left(\alpha_1^2+\gamma_1^2\right)\right]\left[1-\rho_{22}\left(\alpha_2^2+\gamma_2^2\right)\right] - \rho_{12}\rho_{21}\left(\alpha_1^2+\gamma_1^2\right)\left(\alpha_2^2+\gamma_2^2\right)}$$

B.9 Regime Switching ARCH(1) Model - Non-Centered

Definition

Let S_t denote the regime at time t with transition matrix $P = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}$.

$X(t) = \mu_{S_t} + \alpha_{S_t}\left[X(t-1) - \mu_{S_{t-1}}\right] + \epsilon(t)$ where $\epsilon(t) \sim N\left(0, \sigma^2_{S_t}(t)\right)$ and $\sigma^2_{S_t}(t) = \lambda^2_{S_t} + \gamma^2_{S_t} X^2(t-1)$

B.9.1 Stationary Mean and Stationary Variance

For details of derivation, see Section B.7.1.

$$E[X(t)] = \pi_1\mu_1 + \pi_2\mu_2$$

$$Var(X(t)) = \pi_1 V_1 + \pi_2 V_2 + (\mu_1 - \mu_2)^2 \pi_1\pi_2$$

where

$$\begin{cases} V_1 = \frac{\pi_1\rho_{11}}{\pi_1}\left[\left(\alpha_1^2 + \gamma_1^2\right)V_1 + \lambda_1^2 + \gamma_1^2\mu_1^2\right] + \frac{\pi_2\rho_{21}}{\pi_1}\left[\left(\alpha_1^2 + \gamma_1^2\right)V_2 + \lambda_1^2 + \gamma_1^2\mu_2^2\right] \\ V_2 = \frac{\pi_1\rho_{12}}{\pi_2}\left[\left(\alpha_2^2 + \gamma_2^2\right)V_1 + \lambda_2^2 + \gamma_2^2\mu_1^2\right] + \frac{\pi_2\rho_{22}}{\pi_2}\left[\left(\alpha_2^2 + \gamma_2^2\right)V_2 + \lambda_2^2 + \gamma_2^2\mu_2^2\right] \end{cases}$$

Solving the system of equations, one gets

$$V_1 = \frac{\left[1 - \rho_{22}\left(\alpha_2^2 + \gamma_2^2\right)\right]\left[\lambda_1^2 + \alpha_1^2\left(\rho_{11}\mu_1^2 + \rho_{12}\mu_2^2\right)\right]}{\left[1 - \rho_{11}\left(\alpha_1^2 + \gamma_1^2\right)\right]\left[1 - \rho_{22}\left(\alpha_2^2 + \gamma_2^2\right)\right] - \rho_{12}\rho_{21}\left(\alpha_1^2 + \gamma_1^2\right)\left(\alpha_2^2 + \gamma_2^2\right)} + \frac{\rho_{12}\left(\alpha_1^2 + \gamma_1^2\right)\left[\lambda_2^2 + \alpha_2^2\left(\rho_{12}\mu_1^2 + \rho_{22}\mu_2^2\right)\right]}{\left[1 - \rho_{11}\left(\alpha_1^2 + \gamma_1^2\right)\right]\left[1 - \rho_{22}\left(\alpha_2^2 + \gamma_2^2\right)\right] - \rho_{12}\rho_{21}\left(\alpha_1^2 + \gamma_1^2\right)\left(\alpha_2^2 + \gamma_2^2\right)}$$

$$V_2 = \frac{\rho_{21}\left(\alpha_2^2 + \gamma_2^2\right)\left[\lambda_1^2 + \alpha_1^2\left(\rho_{11}\mu_1^2 + \rho_{12}\mu_2^2\right)\right]}{\left[1 - \rho_{11}\left(\alpha_1^2 + \gamma_1^2\right)\right]\left[1 - \rho_{22}\left(\alpha_2^2 + \gamma_2^2\right)\right] - \rho_{12}\rho_{21}\left(\alpha_1^2 + \gamma_1^2\right)\left(\alpha_2^2 + \gamma_2^2\right)} + \frac{\left[1 - \rho_{11}\left(\alpha_1^2 + \gamma_1^2\right)\right]\left[\lambda_2^2 + \alpha_2^2\left(\rho_{21}\mu_1^2 + \rho_{22}\mu_2^2\right)\right]}{\left[1 - \rho_{11}\left(\alpha_1^2 + \gamma_1^2\right)\right]\left[1 - \rho_{22}\left(\alpha_2^2 + \gamma_2^2\right)\right] - \rho_{12}\rho_{21}\left(\alpha_1^2 + \gamma_1^2\right)\left(\alpha_2^2 + \gamma_2^2\right)}$$

B.10 Regime Switching ARCH(1) Model - Proportional

Definition

Let S_t denote the regime at time t with transition matrix $P = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}$.

$X(t) = \mu_{S_t} + \alpha_{S_t}\left[X(t-1) - \mu_{S_{t-1}}\right] + \epsilon(t)$ where $\epsilon(t) \sim N\left(0, \sigma^2_{S_t}(t)\right)$ and $\sigma^2_{S_t}(t) = \gamma^2_{S_t} X^2(t-1)$

B.10.1 Stationary Mean and Stationary Variance

For details of derivation, see Section B.9.1. with $\lambda_1 = \lambda_2 = 0$.

$E[X(t)] = \pi_1\mu_1 + \pi_2\mu_2$

$Var(X(t)) = \pi_1 V_1 + \pi_2 V_2 + (\mu_1 - \mu_2)^2 \pi_1\pi_2$

where

$$V_1 = \frac{\left[1 - \rho_{22}\left(\alpha_2^2 + \gamma_2^2\right)\right]\alpha_1^2\left(\rho_{11}\mu_1^2 + \rho_{12}\mu_2^2\right)}{\left[1 - \rho_{11}\left(\alpha_1^2 + \gamma_1^2\right)\right]\left[1 - \rho_{22}\left(\alpha_2^2 + \gamma_2^2\right)\right] - \rho_{12}\rho_{21}\left(\alpha_1^2 + \gamma_1^2\right)\left(\alpha_2^2 + \gamma_2^2\right)} + \frac{\rho_{12}\left(\alpha_1^2 + \gamma_1^2\right)\alpha_2^2\left(\rho_{12}\mu_1^2 + \rho_{22}\mu_2^2\right)}{\left[1 - \rho_{11}\left(\alpha_1^2 + \gamma_1^2\right)\right]\left[1 - \rho_{22}\left(\alpha_2^2 + \gamma_2^2\right)\right] - \rho_{12}\rho_{21}\left(\alpha_1^2 + \gamma_1^2\right)\left(\alpha_2^2 + \gamma_2^2\right)}$$

$$V_2 = \frac{\rho_{21}\left(\alpha_2^2 + \gamma_2^2\right)\alpha_1^2\left(\rho_{11}\mu_1^2 + \rho_{12}\mu_2^2\right)}{\left[1 - \rho_{11}\left(\alpha_1^2 + \gamma_1^2\right)\right]\left[1 - \rho_{22}\left(\alpha_2^2 + \gamma_2^2\right)\right] - \rho_{12}\rho_{21}\left(\alpha_1^2 + \gamma_1^2\right)\left(\alpha_2^2 + \gamma_2^2\right)} + \frac{\left[1 - \rho_{11}\left(\alpha_1^2 + \gamma_1^2\right)\right]\alpha_2^2\left(\rho_{21}\mu_1^2 + \rho_{22}\mu_2^2\right)}{\left[1 - \rho_{11}\left(\alpha_1^2 + \gamma_1^2\right)\right]\left[1 - \rho_{22}\left(\alpha_2^2 + \gamma_2^2\right)\right] - \rho_{12}\rho_{21}\left(\alpha_1^2 + \gamma_1^2\right)\left(\alpha_2^2 + \gamma_2^2\right)}$$

B.11 Transfer Function Model - Two Lags

Definition

$X(t) = \beta_1 Y(t) + \beta_2 Y(t-1) + e(t)$ where

$e(t) = \mu + \alpha[e(t-1) - \mu] + \epsilon(t)$ and $\epsilon(t) \sim N\left(0, \sigma^2\right)$

where $Y(t)$ follows the RSARCH(P) model (see Section B.10).

Let $Z(t) = \beta_1 Y(t) + \beta_2 Y(t-1) + \mu$

Then $X(t) - Z(t) = \alpha[X(t-1) - Z(t-1)] + \epsilon(t)$

B.11.1 Log-likelihood Function

By conditioning on $e(0)$, one gets

$$l(\boldsymbol{\theta}) = -(n-1)ln\sigma - \frac{n-1}{2}ln(2\pi) - \frac{1}{2\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]^2$$

B.11.2 First Derivative of the Log-likelihood Function

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} = \frac{1-\alpha}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} = \frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)][x(t-1) - z(t-1)]$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \sigma} = -\frac{n-1}{\sigma} + \frac{1}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]^2$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \beta_1} = \frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)][y(t) - \alpha y(t-1)]$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \beta_2} = \frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)][y(t-1) - \alpha y(t-2)]$$

B.11.3 Second Derivative of the Log-likelihood Function

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu^2} = -\frac{(n-1)(1-\alpha)^2}{\sigma^2}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \alpha} = -\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)] - \frac{1-\alpha}{\sigma^2}\sum_{t=1}^{n-1}[x(t-1) - z(t-1)]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\mu\partial\sigma} = -\frac{2(1-\alpha)}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\mu\partial\beta_1} = -\frac{1-\alpha}{\sigma^2}\sum_{t=1}^{n-1}[y(t) - \alpha y(t-1)]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\mu\partial\beta_2} = -\frac{1-\alpha}{\sigma^2}\sum_{t=1}^{n-1}[y(t-1) - \alpha y(t-2)]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha^2} = -\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t-1) - z(t-1)]^2$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha\partial\sigma} = -\frac{2}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)][x(t-1) - z(t-1)]$$

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha\partial\beta_1} = &-\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t-1) - z(t-1)][y(t) - \alpha y(t-1)] \\ &-\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]y(t-1)\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha\partial\beta_2} = &-\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t-1) - z(t-1)][y(t-1) - \alpha y(t-2)] \\ &-\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]y(t-2)\end{aligned}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\sigma^2} = \frac{n-1}{\sigma^2} - \frac{3}{\sigma^4}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]^2$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\sigma\partial\beta_1} = -\frac{2}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)][y(t) - \alpha y(t-1)]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\sigma\partial\beta_2} = -\frac{2}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)][y(t-1) - \alpha y(t-2)]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta_1^2} = -\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[y(t) - \alpha y(t-1)]^2$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta_1\partial\beta_2} = -\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[y(t) - \alpha y(t-1)][y(t-1) - \alpha y(t-2)]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta_2^2} = -\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[y(t-1) - \alpha y(t-2)]^2$$

B.11.4 Stationary Mean and Stationary Variance

$$E[X(t)] = (\beta_1 + \beta_2)\, E[Y(t)] + \mu$$

$$Var(X(t)) = V(Y) + \frac{\sigma^2}{1-\alpha^2}$$

where

$$V(Y) = \sum_{s_t=1}^{2}\sum_{s_{t-1}=1}^{2}\left[\beta_1^2 V_{s_t}(Y) + \beta_2^2 V_{s_{t-1}}(Y) + 2\beta_1\beta_2 C_{s_{t-1},s_t}(Y)\right]\pi_{s_{t-1}}\rho_{s_{t-1},s_t}$$
$$+ \sum_{s_t=1}^{2}\sum_{s_{t-1}=1}^{2}\left[\mu_{s_{t-1},s_t}(Y) - \bar{\mu}(Y)\right]^2\pi_{s_{t-1}}\rho_{s_{t-1},s_t}$$

and

$s_t =$ regime at time t

$\mu_{s_t}(Y) = E\left[Y(t)\,|s_t\right]$

$V_{s_t}(Y) = Var\left(Y(t)\,|s_t\right)$ (see section B.10.1)

$C_{s_{t-1},s_t}(Y) = Cov\left(Y(t), Y(t-1)\,|s_t, s_{t-1}\right) = \alpha_{s_t} V_{s_{t-1}}(Y)$

$\mu_{s_{t-1},s_t}(Y) = \beta_1\mu_{s_t}(Y) + \beta_2\mu_{s_{t-1}}(Y)$

$$\bar{\mu}(Y) = \sum_{s_t=1}^{2}\sum_{s_{t-1}=1}^{2}\mu_{s_{t-1},s_t}(Y)\pi_{s_{t-1}}\rho_{s_{t-1},s_t}$$

B.12 Transfer Function Model - One Lag

Definition

$X(t) = \beta_1 Y(t) + e(t)$ where

$e(t) = \mu + \alpha[e(t-1) - \mu] + \epsilon(t)$ and $\epsilon(t) \sim N\left(0, \sigma^2\right)$

where $Y(t)$ follows the RSARCH(P) model (see Section B.10).

Let $Z(t) = \beta_1 Y(t) + \mu$

Then $X(t) - Z(t) = \alpha[X(t-1) - Z(t-1)] + \epsilon(t)$

B.12.1 Log-likelihood Function

By conditioning on $e(0)$, one gets

$$\begin{aligned} l(\boldsymbol{\theta}) = & -(n-1)ln\sigma - \frac{n-1}{2}ln(2\pi) \\ & - \frac{1}{2\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]^2 \end{aligned}$$

B.12.2 First Derivative of the Log-likelihood Function

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} = \frac{1-\alpha}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} = \frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)][x(t-1) - z(t-1)]$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \sigma} = -\frac{n-1}{\sigma} + \frac{1}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]^2$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \beta_1} = \frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)][y(t) - \alpha y(t-1)]$$

B.12.3 Second Derivative of the Log-likelihood Function

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu^2} = -\frac{(n-1)(1-\alpha)^2}{\sigma^2}$$

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \alpha} = &-\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)] \\ &-\frac{1-\alpha}{\sigma^2}\sum_{t=1}^{n-1}[x(t-1) - z(t-1)]\end{aligned}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \sigma} = -\frac{2(1-\alpha)}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu \partial \beta_1} = -\frac{1-\alpha}{\sigma^2}\sum_{t=1}^{n-1}[y(t) - \alpha y(t-1)]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha^2} = -\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t-1) - z(t-1)]^2$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \sigma} = -\frac{2}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)][x(t-1) - z(t-1)]$$

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \beta_1} = &-\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t-1) - z(t-1)][y(t) - \alpha y(t-1)] \\ &-\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]y(t-1)\end{aligned}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma^2} = \frac{n-1}{\sigma^2} - \frac{3}{\sigma^4}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)]^2$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \sigma \partial \beta_1} = -\frac{2}{\sigma^3}\sum_{t=1}^{n-1}[x(t) - z(t) - \alpha x(t-1) + \alpha z(t-1)][y(t) - \alpha y(t-1)]$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_1^2} = -\frac{1}{\sigma^2}\sum_{t=1}^{n-1}[y(t) - \alpha y(t-1)]^2$$

B.12.4 Stationary Mean and Stationary Variance

$E[X(t)] = \beta_1 E[Y(t)] + \mu$

$Var(X(t)) = \beta_1^2 Var(Y(t)) + \dfrac{\sigma^2}{1-\alpha^2}$

Bibliography

[1] Aitken, W. H. (1996). *A problem-solving approach to pension funding and valuation* (2nd ed.). Winsted, Conn. : ACTEX Publications.

[2] Akaike, H. (1974). A new look at the statistical model identification. *Transactions on Automatic Control*, 19(6), 716-723.

[3] Aon Consulting. (2008). Replacement ratio study: A measurement tool for retirement planning. Retrieved from http://www.aon.com/about-aon/intellectual-capital/attachments/human-capital-consulting/RRStudy070308.pdf

[4] Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control* (2nd ed.). San Francisco: Holden-Day.

[5] Cairns, A. J. G. (1996). Continuous-time pension-fund modelling. *Proceedings of the 6th AFIR Colloquium* (Vol. 1, pp. 609-624), Nurnberg, October 1996.

[6] Cairns, A. J. G., & Parker, G. (1997). Stochastic pension fund modelling. *Insurance: Mathematics and Economics*, 21, 43-79.

[7] Chang, S. C., & Cheng, H. Y. (2002). Pension valuation under uncertainties: Implementation of a stochastic and dynamic monitoring system. *The Journal of Risk and Insurance*, 69(2), 171-192.

[8] Cochrane J. H. (1994). Permanent and transitory components of GNP and stock prices. *The Quarterly Journal of Economics*, 109(1), 241-265.

[9] Cox J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2), 385-407.

[10] Diamond, P. A. (1977). A framework for social security analysis. *Journal of Public Economics*, 8(3), 275-298.

[11] Dufresne, D. (1989). Stability of pension systems when rates of return are random. *Insurance: Mathematics and Economics*, 8, 71-76.

[12] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1007.

[13] Haberman, S. (1994). Autoregressive rates of return and the variability of pension contributions and fund levels for a defined benefit pension scheme. *Insurance: Mathematics and Economics*, 14, 219-240.

[14] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationarity time series and the business cycle. *Econometrica*, 57, 357-384.

[15] Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64, 307-333.

[16] Hardy, M. (2003). *Investment guarantees: Modeling and risk management for equity-linked life insurance*. US: John Wiley & Sons, Inc.

[17] Inflation-control target. (n.d.). *Bank of Canada*. Retrieved from http://www.bankofcanada.ca/rates/indicators/key-variables/inflation-control-target/

[18] Key interest rate: Target for the overnight rate. (n.d.). *Bank of Canada*. Retrieved from http://www.bankofcanada.ca/monetary-policy-introduction/key-interest-rate/

[19] Koop, G., Milas, C. Milas, & Osborn, D. (2008). Regime-switching models in economics and finance. *Studies in Nonlinear Dynamics & Econometrics*, 12(3). Retrieved from http://www.bepress.com/snde/vol12/iss3/

[20] Özmucur, S. (2008). Wage and productivity differentials in private and public manufacturing: The case of turkey. Retrieved from http://129.3.20.41/eps/io/papers/0303/0303002.pdf

[21] Simms, D. (2011, July 12). Why Canadians should care about a U.S. debt default. *CBC News*. Retrieved from http://www.cbc.ca/news/business/story/2011/07/12/f-us-default-canadian-effects.html

[22] Thury, G. (1979). Causality relationships between macroeconomic variables: An empirical investigation for Austria. *Empirica*, 6(2), 205-216.

[23] Wilkie, A. D. (1986). A stochastic investment model for actuarial use. *Transactions of the Faculty of Actuaries*, 39, 341-403.

[24] Wilkie, A. D. (1995). More on a stochastic asset model for actuarial use. *British Actuarial Journal*, 1(5), 777-964.

[25] Zhang C. B., & Hou, R. J. (2011). Optimal portfolios for DC pension with stochastic salary. *2011 International Conference on Business Management and Electronic Information* (Vol. 1, pp. 589-592), Guangzhou, October 2011.