A Pension Fund Modelling Study: CCland

Executive Summary

The goal of this task is to build a stochastic investment model for future asset and liability purposes. To achieve this, we need to make assumptions in the early stage. First, it is assumed that we only have limited given economic variables. Then the ARIMA model is chosen to be fitted into time series after multiple calculations and trials, and as to the residuals, it is assumed to follow the white noise process (i.e. normal distribution).

In the next stage data analysis, after transforming the indexes into the rate of return, as to each economic variable, the time series is formed by the quarterly data from the second quarter of 1980 to the second quarter of 2016. Then we do Augmented Dickey-Fuller (ADF) and Phillips–Perron (PP) tests on the time series to ensure that it is a non-seasonal and stationary series.

Then in the modeling part, Wilkie's (1986,1995) structure is employed in this report due to well-appropriate fitting and similar background information. In this structure, price inflation is assumed to be a driving factor, and the overall cascade system is a price-inflation- driving system. Price inflation is independent due to driving factors. As to the other economic variables, we do separate modelings for each economic variable and analyze the relationship between them to select the best-fitted model by using the tests. Finally, we use the Jarque-Bera test to do model validation by checking the normality and correlation of residuals.

After getting revised formulae for economic variables, the future investment returns are projected by running the model over 5 years (20 quarters). Three different sets of asset allocation ratios are chosen to show the sensitivity of equity and bond in respective of total future returns.

1 Introduction

This essay illustrates the structure of the model for future asset and liability. It is assumed that there are only six economic variables in this model. We make use of original data of each economic variable and analyze the time series of them. After getting the original formula for each of them, we imply the Wilkie structure to discuss their relationship and use the tests to check whether the model is reasonable and appropriate. Finally, we use revised formulae to forecast future trends of the economic variable and then project future investment returns by combing six economic variables.

1.1 Assumptions

In this report, we make a forecast based on the model we build given data of the past several decades. The basic premise is that the pattern of asset classes in the past can be used to project claims into the future. In other words, the performance of economic variables in this report is expected to develop over the next 5 years in the same way as they have in the past. The variables involved in our model are price inflation rate, wage inflation rate, equity yield, equity dividend yield, long-term interest rate, and short-term interest rate.

We then apply the price-inflation-driving system based on available data to build economic models and use these models to forecast future asset returns. The system is mainly referred from the Wilkie-type structure and is widely accepted in the UK. We also assume that the economic assumptions and variables are similar to the situation in Australia so they can be compatible with the Wilkie-type model.

Although the proposed models are reasonable and relatively consistent with reality, the real-life economic variables are vulnerable to many other external factors that we do not discuss below. Therefore, the prediction is better used for short-term investment projections only and should be modified frequently.

2 Data

We analyze quarterly data for modeling purposes from the second quarter of 1980 to the second quarter of 2016. It is the largest common period for all available data. Then we convert all available data into a consistent and appropriate period, the nominal compound quarterly rate of change, for modeling purpose. It is known that there are 6 categories of data, that is CPI index, wage income, equity index dividend, equity price index, 10-year government bond yields, and 3-month treasury bill rates. As to CPI, Wage Income, Equity Price Index, the nominal compound quarterly rate is the rate of change in percentage

$$\frac{I(t+1)}{I(t)} - 1,$$

where $I(t)$ is the index I in corresponding time t, and we only use data of March, June, September, and December to derive the nominal quarterly rate? As to Equity Index Dividend Yield, 10 Year Government Bond Yield and 3-month Treasury Bill rate, nominal compound quarterly rates are simply extracted from the raw rates in March, June, September, and December for each quarter.

Then we use Augmented Dickey-Fuller (ADF) and Phillips–Perron (PP) tests for seasonal unit root testing. The main goal of these two tests is to decide and get a stationary and non-seasonal series by differencing time series with one unit root or seasonal unit root.

We denote variables for the time of quarters t by lower case letters. For instance, $p(t)$ is used for the price inflation rate, and $w(t)$ is used for the wage inflation rate. Also, we use to denote the seasonal difference for variables.

Price inflation rate: We use the nominal compound percentage change of the consumer price index to reflect the price inflation rate. The results of the ADF and PP test show seasonality and non-stationarity in our derived price inflation rate series $\{p(t)\}$. So we remove the tested features by taking the first-order difference $\Delta p(t)$ and do some tests on the new series. Figure 1 presents both the original and adjusted series.

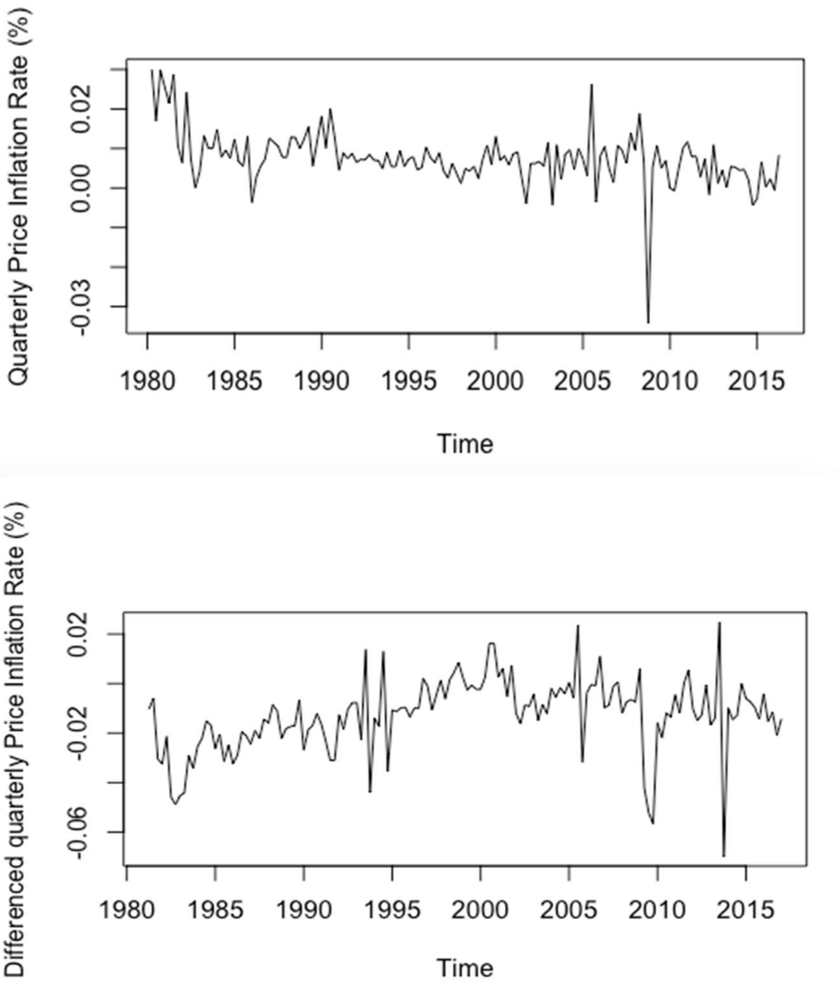
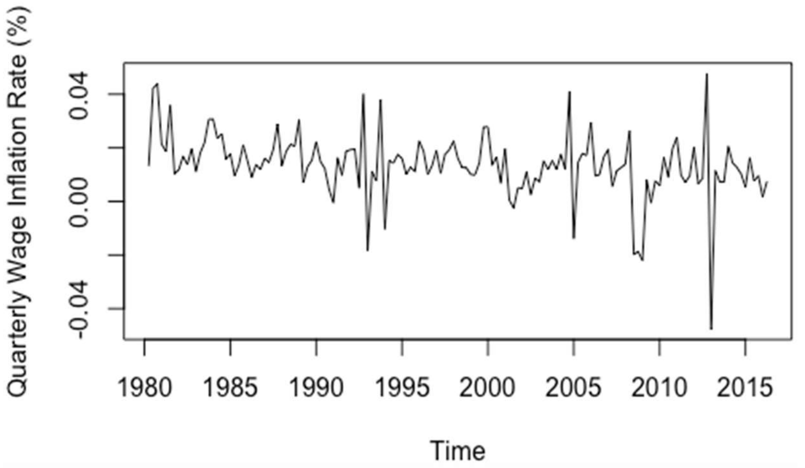


Figure 1

Wage inflation rate: We compute the wage inflation rate of change given the monthly wage income data. The plot of the wage inflation rate is shown in Figure 2. The strong increasing trend is because of the considerable development of economics and society from 1980. We stabilize such powerful movement by seasonal differencing the data, and the new series is also plot in the second graph of Fig. 2. Both ADF and PP unit root tests of the seasonal differenced wage inflation rate series suggest stationarity as well as non-seasonality.



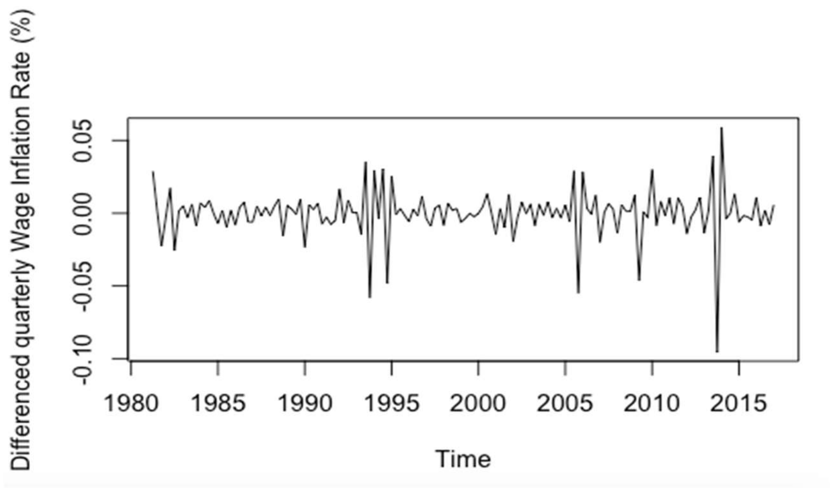


Figure 2

Equity Yield: The quarterly equity yield series is derived from the equity price index. The results of unit root tests suggest seasonality. Thus, we do the same process and produce seasonal differenced stationary series with no seasonality. The original series and seasonal differenced equity yield plots are presented in Figure 3.

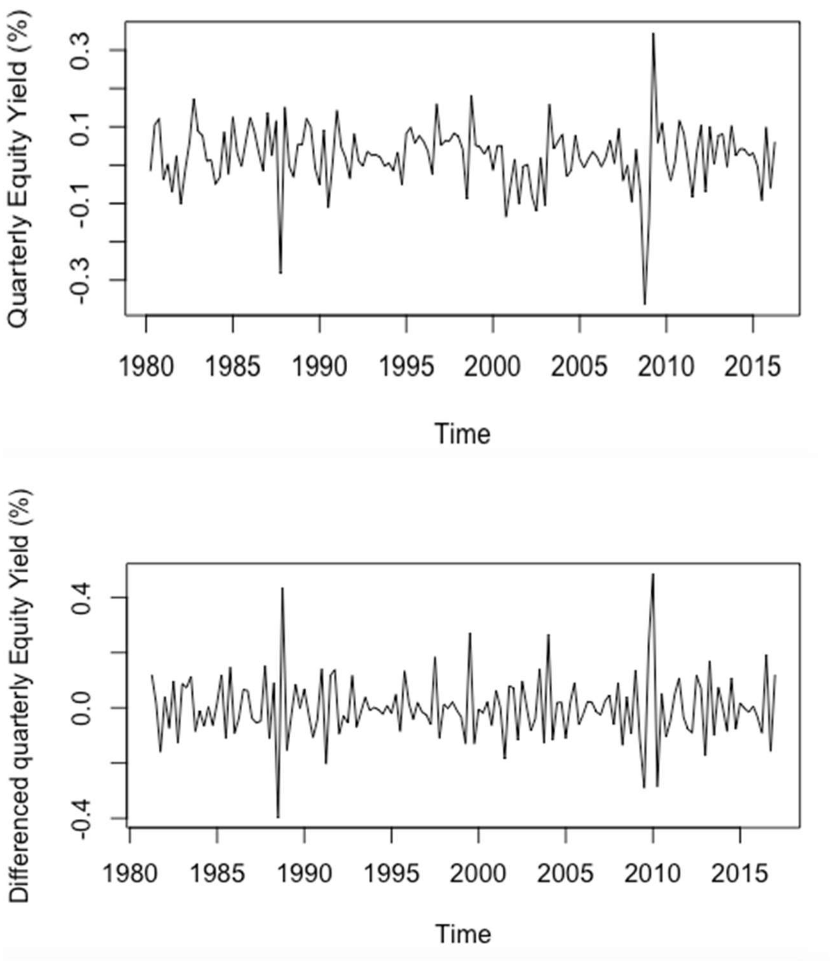


Figure 3.

Equity Dividend Yield: We produce nominal equity dividend yield series compounded quarterly from nominal yield series compounded monthly. The plot of quarterly equity dividend yield shows an overall non-stationary decreasing trend, so we produce differenced quarterly series. Both plots are presented in Figure 4. The ADF and PP tests demonstrate stationary and non-seasonality.

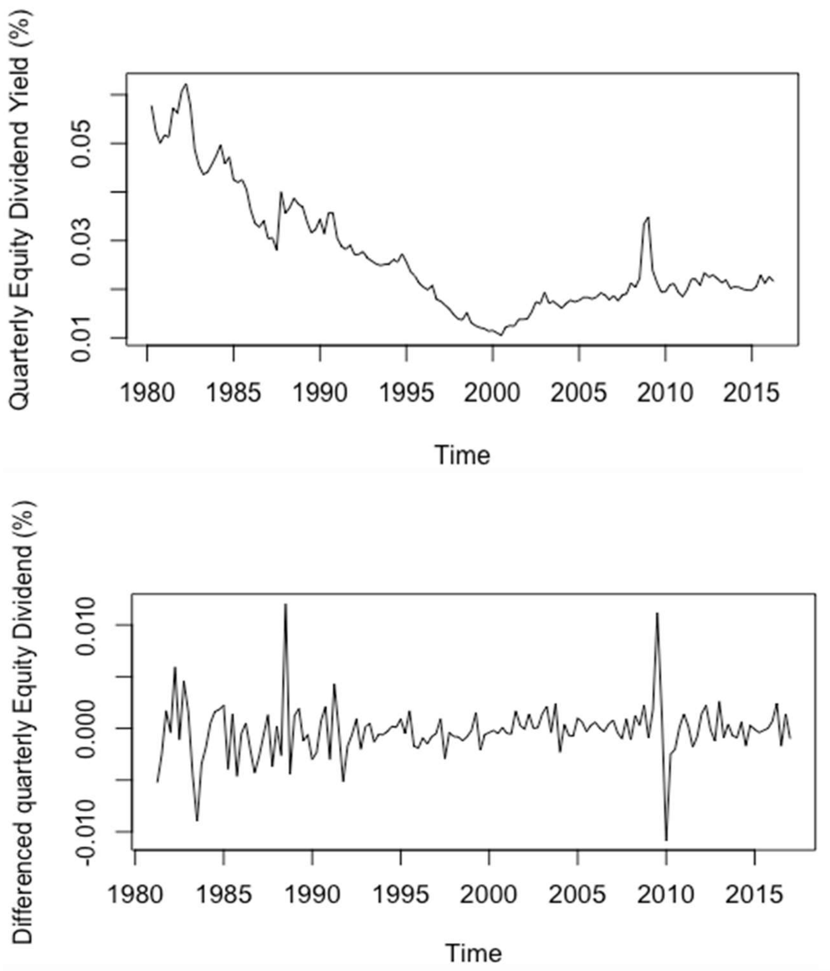


Figure 4.

Long-term interest rate: We use the 10-year government bond yield to represent the long-term interest rate. Both unit root tests of the long-term interest rate series show one unit root. So, we difference the data to obtain a stationary sequence. Similarly, the plots for the original and revised series are in Figure 5.

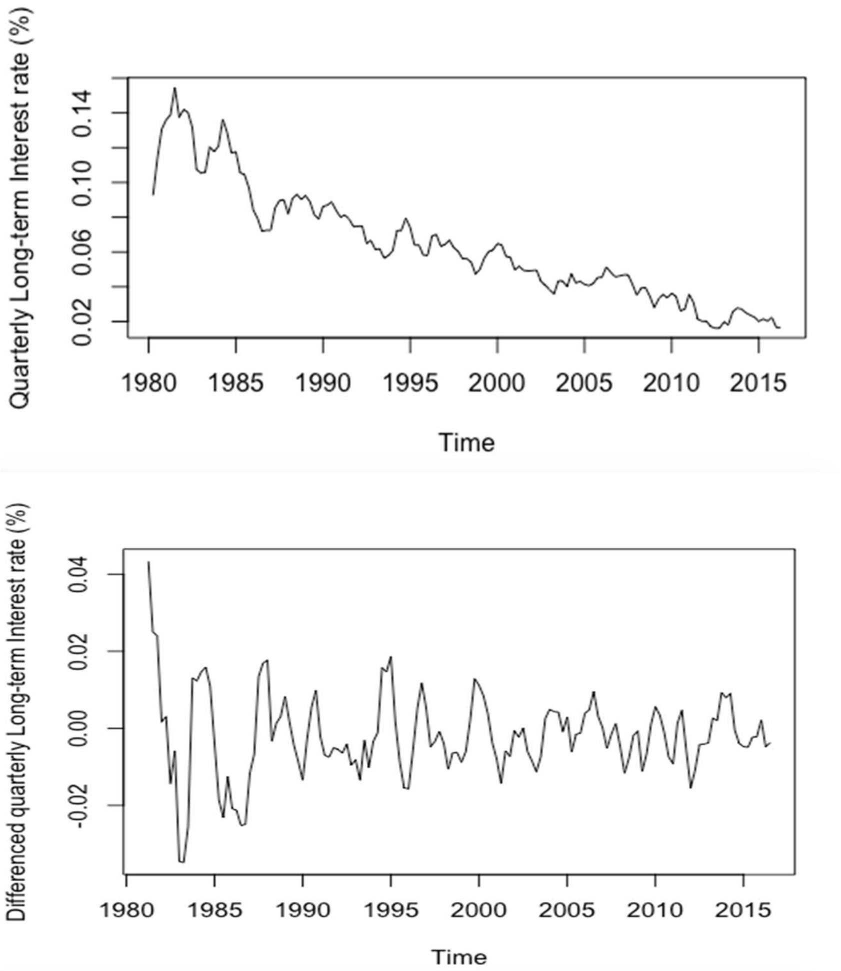


Figure 5.

Short-term interest rate: We use the 3-month Treasury bills rate to describe short-term interest rate {s(t)}. The unit root tests suggest that there exists a unit root in the original time series. We again remove the non-stationarity by producing the seasonal differenced series {Δs(t)}. Both plots are shown in Figure 6.

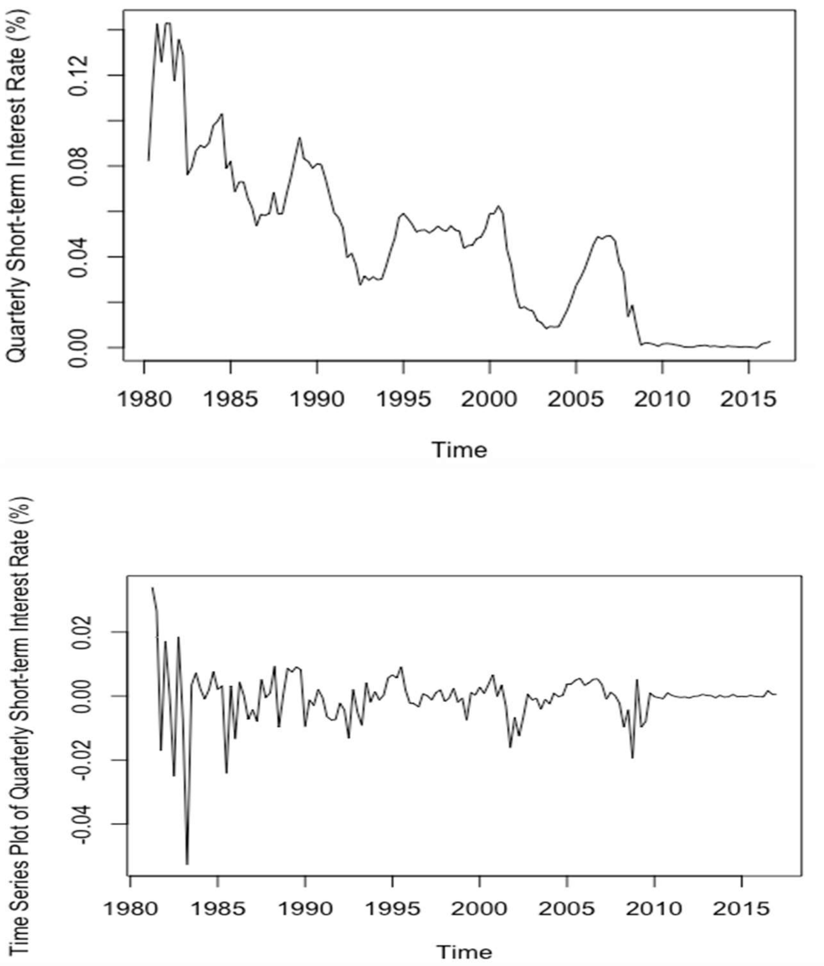


Figure 6.

All variables used in this report are summarized in Table 1.

Table 1. Variables.

Variable	nominal compounded	Stationary
Price inflation rate	$p(t)$	$\Delta p(t)$
Wage inflation rate	$w(t)$	$\Delta w(t)$
Equity yield	$ey(t)$	$\Delta ey(t)$
Equity Dividend yield	$ed(t)$	$\Delta ed(t)$
Long-term interest rate	$l(t)$	$\Delta l(t)$
Short-term interest rate	$s(t)$	$\Delta s(t)$

3. Models

3.1 Model identification

We believe that it is hardly reasonable to assume that the returns of bonds and equities are independent and unrelated to other market factors. So, we use the univariate model only for factors that are assumed independent. In this report, we employ the Wilkie (1986, 1995) structure that is widely accepted in UK and assumes the price inflation rate to be the driving force for all market variables. The cascade structure of our price-inflation-driving system we adopted in this report is as followed in Fig. 7 as a result of model selection derived in the following sections. Greek letters are used to denote coefficients in the models while ε is used to denote error terms with corresponding subscripts.

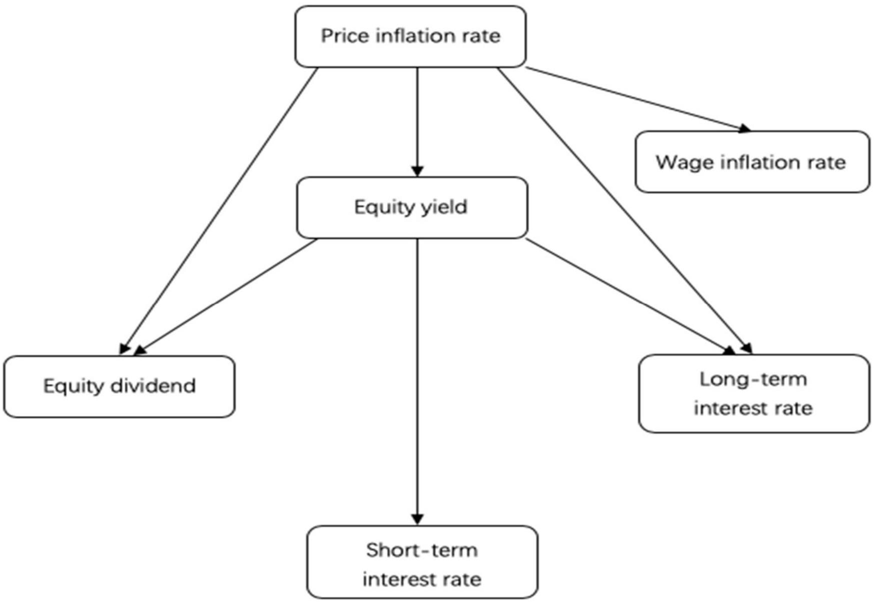


Fig. 7. The structure of the price-inflation-driving system.

Before deciding the overall cascade system, we derive ARIMA models for all variables with the help of autocorrelation function (ACF) and partial autocorrelation function (PACF) to identify lags of variables that exert significant influences. We also test and compare models with different estimated coefficients and variable linkages by Akaike information criterion (AIC) and Bayesian information criterion (BIC) as goodness-of-fit measures to determine the best fit. Smaller values of both AIC and BIC are the sign of better-fitted model.

3.2 Model checking

We have several criterions to check the appropriateness of models. Firstly, we assume

the error terms or residuals of estimated models to follow the white noise process. So, we check normality and correlations of residuals by Jarque-Bera test (with *JB* statistic), where the test statistic should be significantly small to demonstrate normality.
Then, we conduct out-of-sample validation to evaluate the forecasting performance of models. We use observations from quarters of 1980 to 2000 to estimate the model parameters, and then use these parameters to generate a forecast sample for quarters of 2001 to 2016. After that, we compare the actual observations with the forecasted values and generate the root mean square forecast error. Finally, we choose the model with the smallest RMSFE, which all based on non-differenced variables. However, for model fitting, we always use the whole data period from the second quarter of 1980 to the second quarter of 2016.

3.3 Model estimation.

3.3.1 Price inflation rate

This is the driving force in the cascade modeling structure. Therefore we need to build a univariate ARIMA model, whose seasonal difference follows an AR(3) process:

$$\Delta p(t) = p(t) - p(t-1)$$
$$\Delta p(t) = c_1 \Delta p(t-1) + c_2 \Delta p(t-2) + c_3 \Delta p(t-3) + \varepsilon_p$$

Where c_i is the corresponding constant coefficient for $\Delta p(t-i)$.

This model indicates that the seasonal differenced price inflation rate is highly related and determined by the previous three consecutive seasonal differences.

By simply computation we get the final model for price inflation rate $p(t)$ is:

$$p(t) = \alpha_p p(t-1) + \beta_p p(t-2) + \gamma_p p(t-3) + \delta_p p(t-4) + \varepsilon_p$$

The estimations and model checking of results are in Table 2. The residual diagnostic check shows that the normality assumption of error terms are satisfied. And there are no autocorrelations in residuals.

Table 2. Estimations for price inflation rate.

	Price inflation rate
α_p	1.2461
β_p	0.1681
γ_p	- 0.2382
δ_p	-0.176
S.E. of regression	0.000175
Jarque-Bera	< 2.2e-16
AIC	-829.9582

BIC	-830.6801

Significant at the 5% level.

AIC, Akaike information criterion; BIC, Bayesian information criterion。

3.3.2 Wage inflation rate

We have tried different models with different lag orders to determine the relationship between wage inflation and price inflation, and this model turns out to be the best one following from our model selection criterion:

$$\Delta w(t) = w(t) - w(t-1)$$

$$\Delta w(t) = \alpha_w \Delta w(t-1) + \beta_w \Delta p(t) + \varepsilon_w$$

The estimations and model checking of results are in Table 3. The seasonal differenced wage inflation is modelled by a positive proportion of contemporaneous seasonal differenced price inflation, but influenced negatively with previous one-period wage inflation, which is consistent with reality that the good economic performance (as measured by CPI movement) leads to increases in wages, while adjusted with some fluctuations in response for performance of individual businesses. The Jarque–Bera normality test demonstrates the normally distributed residuals. Autocorrelations of residuals are insignificant and there is no evidence of conditional heteroscedasticity.

Table 3. Estimations for wage inflation rate.

	Wage inflation rate
α_w	-0.5826
β_w	0.1245
S.E. of regression	0.0001682
Jarque-Bera	< 2.2e-16
AIC	-838.355
BIC	-837.0769

Significant at the 5% level.

AIC, Akaike information criterion; BIC, Bayesian information criterion.

3.3.3 Equity yield

The seasonal differenced equity yield is linked with the seasonal difference price inflation rate as follows:

$$\Delta ey(t) = ey(t) - ey(t-1)$$

$$\Delta ey(t) = \alpha_{ey} \Delta ey(t-1) + \beta_{ey} \Delta p(t) + \varepsilon_{ey}$$

To interpret the model, the seasonal differenced equity yield is modelled by a proportion β_{ey} of the contemporaneous seasonal differenced price inflation rate together with a proportion α_{ey} of seasonal differenced equity yields of last period. This is intuitive because we expect changes of equity yields to be correlated movement of

CPI and equity itself. The estimated coefficients and diagnostic checking of results are presented in Table 4. We can see that residuals are normally distributed, and there are no autocorrelations in residuals.

Table 4. Estimations for equity yield.

	Equity yield
α_{ey}	-0.4930
β_{ey}	-0.0676
S.E. of regression	0.01029
Jarque-Bera	< 2.2e-16
AIC	-246.0325
BIC	-244.7544

Significant at the 5% level.

AIC, Akaike information criterion; BIC, Bayesian information criterion.

3.3.4 Equity dividend yield

The equity dividend yield is derived to be driven by the contemporaneous price inflation rate and equity yield. The model is:

$$\Delta ed(t) = ed(t) - ed(t-1)$$

$$\Delta ed(t) = \alpha_{ed}\Delta ed(t-1) + \beta_{ed}\Delta p(t) + \Delta ey(t) + \varepsilon_{ed}$$

Given the model estimations illustrated in table 5, we can see that if the CPI and equity yield increase, the contemporaneous equity dividend yield will decrease. This is analogous to investors receiving less dividend as more equities are retained in business. It is also evident in estimation table that residuals are normally distributed and no autocorrelations.

Table 5. Estimations for equity dividend yield.

	Equity dividend yield
α_{ed}	0.3819
β_{ed}	- 0.0034
γ_{ed}	-0.0147
S.E. of regression	2.935e-06
Jarque-Bera	< 2.2e-16
AIC	- 1421.572
BIC	- 1420.294

Significant at the 5% level.

AIC, Akaike information criterion; BIC, Bayesian information criterion.

3.3.5 Long-term interest rate

The best model we derived for long-term interest rate is

$$\Delta l(t) = l(t) - l(t-3)$$

$$\Delta l(t) = \alpha_l \Delta l(t-1) + \beta_l \Delta p(t) + \gamma_l \Delta ey(t) + \varepsilon_l$$

such that the third differenced long-term interest rate is changing with contemporaneous price inflation rate as well as equity yield. Given the parameter estimations illustrated in table 6, we can see that if the price inflation increases and the equity yield decreases, the contemporaneous long-term interest rate will increase. It is also shown below that the residuals are normally distributed and no autocorrelations.

Table 6. Estimations for Long-term interest rate.

	Long-term interest rate
α_l	0.7557
β_l	0.0521
γ_l	-0.0103
S.E. of regression	5.62e-05
Jarque-Bera	< 2.2e-16
AIC	-979.8676
BIC	-977.9504

Significant at the 5% level.

AIC, Akaike information criterion; BIC, Bayesian information criterion.

3.3.6 Short-term interest rate

Short-term interest rate are only driven by contemporary effect of equity yield in our system. And the model we derive for seasonal differenced short-term interest rate is

$$\Delta s(t) \ = s\ (t) - s(t-1)$$
$$\Delta s\ (t) = \alpha_s \Delta s\ (t-1) + \beta_s \Delta ey\ (t-1) + \varepsilon_s$$

Which also added the effect of a first-order autoregressive component. We also test models that have undifferenced or differenced price inflation and long-term interest rate as explanatory variates, but we did not get better results. And we finalized the model estimation and residual diagnostic checking in Table 7, which shows no evidence of non-normality and autocorrelations in residuals.

Table 7. Estimations for Short-term interest rate.

	Short-term interet rate
α_s	0.0219
β_s	0.0050
S.E. of regression	5.825e-05
Jarque-Bera	< 2.2e-16
AIC	-977.6323
BIC	-976.3542

Significant at the 5% level.

AIC, Akaike information criterion; BIC, Bayesian information criterion.

4 Model Forecasting

The pension fund only invests in equities, bonds, and cash. For each component, the return of equity is calculated by the combination of equity yield and equity dividend yield. Long term bond yield is used for the bond return, short term bond yield is used for the cash return. The price inflation rate is included in the equity yield and equity dividend yield. The total nominal compound quarterly rate of this pension fund is the combination of equity return, bond return and cash return with weighs allocation. Therefore, the fund investment return after the 5-year period is the product of (1+each return/4), where each return is total nominal compound quarterly rate and return/4 is an effective rate of return per quarter. For the 5-year future investment return, we need to forecast 20 total nominal compound quarterly rates. The forecast results for all economic variables are based on 100 stimulations.

Table 8. Revised formulae for forecasting.

Revised Formulae
$p(t) = 1.2461p(t-1) + 0.1681p(t-2) - 0.2382p(t-3) - 0.01 + \varepsilon_p$ μ = -0.0002 σ = 0.0002
$w(t) = -0.5826w(t-1) + 0.1245p(t) - 0.1245p(t-1) + \varepsilon_w$ μ = 0.02 σ = 0.005
$ey(t) = 0.507ey(t-1) + 0.493ey(t-2) - 0.0677p(t) + 0.0677p(t-1) + \varepsilon_{ey}$ μ = 0.01 σ = 0.04
$ed(t) = -0.3819ed(t-1) - 0.0147ey(t) + 0.0147ey(t-1) - 0.0034p(t) + \varepsilon_{ed}$ μ = 0.05 σ = 0.01
$l(t) = 0.7557l(t-1) - 0.0103ey(t) + 0.0103ey(t-1) + 0.0512p(t) - 0.0512p(t-1) + \varepsilon_l$ μ = 0.02 σ = 0.015
$s(t) = 1.0219\,s(t-1) - 0.0219\,s(t-2) + 0.005e\,y(t-1) - 0.005e\,y(t-2) + \varepsilon_s$ μ = 0.006 σ = 0.004

Based on the formulae above, the error of the forecast price inflation rate is relatively small after comparing it with the original data range. Realistically, forecast data has certain growth as time goes by. The forecast for the wage inflation rate has similar features. As to equity yield, from the theoretical perspective, the error of equity yield has a relatively large variable. It is because that equity is the asset that has a relatively higher risk with higher risk-reward in return so that it also possible to obtain a positive and negative return. Then in terms of equity dividend yield, in reality, it cannot be negative, but here it has a negative rate in raw data. The equity dividend yield is expected to have a small grow in the future because of increasing CPI rate and Equity yield. Compare to long-term bond yield, short-term bond yield is relatively lower than long-term bond yield, which is realistic in practice. In terms of risk, short-term bond yield is lower than long-term bond yield, since short-term bond yield is calculated as cash rate (a.k.a risk-free rate).

4.1 Asset allocation ratio

Here we chose 3 different sets of asset allocation ratios to predict future returns. To observe the influence of equity on return, the first set has a large weight on equity. Similarly, to observe the influence of bond on return, the second set has a large weight on bond. The third set has a middleweight for three assets.

When assets put a large weight on equity, for example, 90% equity, 5% bond and 5% cash, the portfolio return will present relatively high variability and high return. There exist extreme values on long tails. It has a high possibility to gain a return between 0.023119306 and 0.040161707. It is because equity is the riskiest asset among three assets in the portfolio, and a large risk premium always is used to compensate investor who bears more risks.

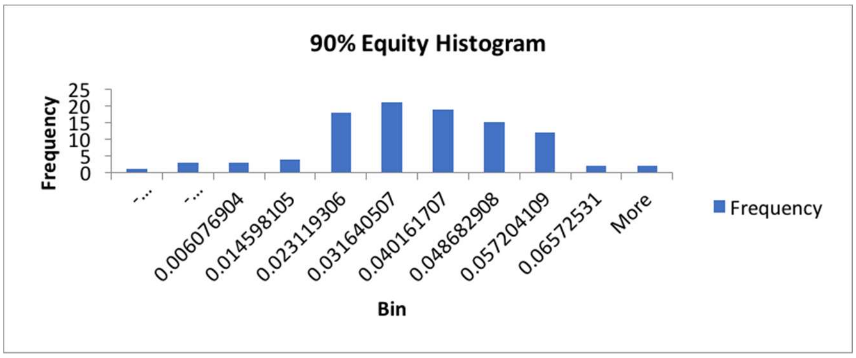


Figure 8.

When assets put a large weight on bond, for example, 90% bond, 5% equity, and 5% cash, the portfolio return will present relatively low variability and low return. It is because the bond is less risky than equity and also lower risk premium return. From the 90% Bond Histogram below, the distribution of return concentrate between 0.015990216 and 0.02128467, there is a relatively lower possibility of occurring extreme values.

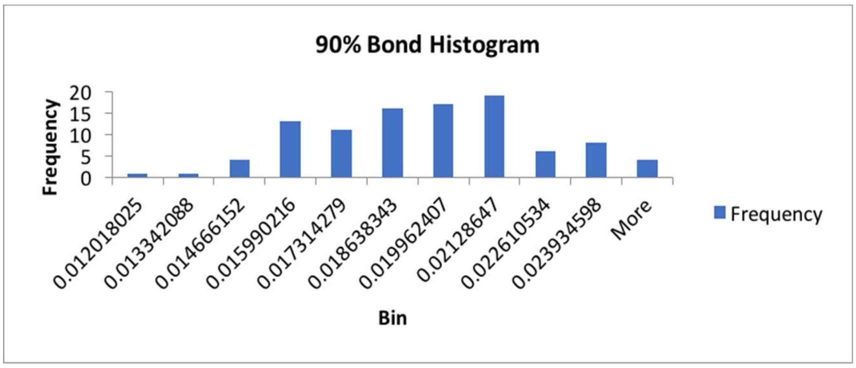


Figure 9.

Finally, when we take the middle weights for three assets, for example, 40% equity, 30% bond and 30% cash, the variability of portfolio investment is between previous two asset allocations (0.007657), and the concentration of return is around 0.01976 and 0.0349 which is higher than 90% Bond but lower than 90% Equity.

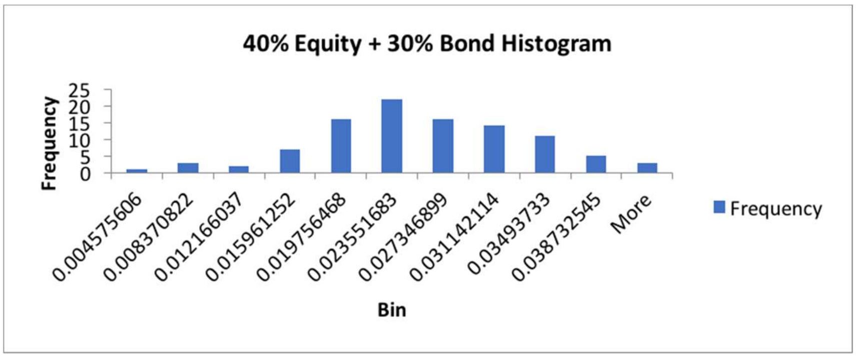


Figure 10.

After analyzing three situations, it can be concluded that a risky asset allocation ratio is likely to bring out high future investment returns and a more volatile return. As to risk-averse investors, they may prefer the asset allocation ratio with high bond weights for lower volatility, for example, 90% bond. As to risk seeker, they may prefer 90% equity asset allocation, since they have the chance to pursue extreme high return.

Reference

Huang, F., Butt, A. and Ho, K. (2014). Stochastic economic models for actuarial use: an example from China. *Annals of Actuarial Science*, 8(2), pp.374-403.

Wilkie, A. (1986). A stochastic investment model for actuarial use. Transactions of Faculty of Actuaries, 39, 341–403. Wilkie, A. (1995). More on a stochastic asset model for actuarial use. *British Actuarial Journal*, 1, 777–964.