Dynamic Retirement Financial Planning Model

Qintian Sun
University of Connecticut - Storrs, qintian.sun@uconn.edu

Follow this and additional works at: https://opencommons.uconn.edu/dissertations

Recommended Citation
Sun, Qintian, "Dynamic Retirement Financial Planning Model" (2019). Doctoral Dissertations. 2207.
https://opencommons.uconn.edu/dissertations/2207
Dynamic Retirement Financial Planning Model

Qintian Sun, Ph.D.
University of Connecticut, 2019

ABSTRACT

We developed a retirement financial planning strategy based on Markov chain modeling of retirement health conditions and Geometric Brownian Motion modeling of asset values. The annual living expenses of a retiree are modeled as basic expenses plus discretionary expenses. Our goal is to solve for the maximum discretionary expenses while healthy, which was obtained using a closed-form solution and quantile optimization technique. The highlight of this model is the use of Kalman Filter for annual recalibration. It allows the model to automatically adjust the suggested amount of discretionary expenses by looking at daily fund values from previous year. After running a lot of simulations and testings, we showed that our dynamic model beats other static models and a naive recalibration model in the sense that it virtually eliminates ruin and is able to let a retiree withdraw the largest possible amount.
Dynamic Retirement Financial Planning Model

Qintian Sun

M.S. University of Connecticut [2016]
B.S. Beijing Normal University [2014]

A Dissertation
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
at the
University of Connecticut

2019
APPROVAL PAGE

Doctor of Philosophy Dissertation

Dynamic Retirement Financial Planning Model

Presented by
Qintian Sun, B.S. Math., M.S. Math.

Major Advisor
__________________________
Jeyaraj Vadiveloo

Associate Advisor
__________________________
Bin Zou

Associate Advisor
__________________________
Guojun Gan

University of Connecticut

2019
# Contents

Ch. 1. Introduction ............................................. 1

Ch. 2. Literature Review ......................................... 3

Ch. 3. Methodology .............................................. 6
   3.1 Notations .................................................. 6
      3.1.1 Health conditions and Annual Expenses ............. 6
      3.1.2 Assets and Annuities .................................. 7
      3.1.3 Kalman Filter .......................................... 8
   3.2 Health Modeling and Spending patterns .................. 8
      3.2.1 Healthy and Unhealthy Life Expectancy ............. 8
      3.2.2 Basic Expenses and Discretionary Expenses .......... 9
      3.2.3 Annuities: SPIA and DIA ............................ 10
   3.3 Optimal Solution .......................................... 11
      3.3.1 Asset Returns Modeling ............................... 11
      3.3.2 Quantile Optimization with Closed-Form Solution .... 13
      3.3.3 Multiple Assets: Allocation and Covariance .......... 17
   3.4 Kalman Filter ............................................. 19
      3.4.1 Derivation and Properties ............................ 19
      3.4.2 Kalman Filter Set Up .................................. 26

Ch. 4. Model Results ............................................ 31
   4.1 Model Summary ............................................ 31
   4.2 Simulation Method ........................................ 33
   4.3 Case Studies .............................................. 34
      4.3.1 Model 1: Fixed Death Time ............................ 37
      4.3.2 Model 2: Random Death Time .......................... 38
      4.3.3 Model 3: HLE and ULE ............................... 38
4.3.4 Model 4: Naive Recalibration .......................................... 39
4.3.5 Model 5: Kalman Filter Recalibration ............................... 43
4.3.6 Model Comparison ....................................................... 44
4.4 More Applications .......................................................... 50
4.4.1 Bequest Needs ............................................................. 50
4.4.2 Liability Cash flow for Fund Manager .............................. 50

Bibliography ..... 53
Chapter 1

Introduction

This research is inspired by the quantile optimization method in Xu 2018 [10] and keeps focusing on the method to help retirees determine the optimal withdrawals from their retirement funds.

We first model the health statuses using Markov chains, which can naturally provide us specific lengths of HLE (Healthy Life Expectancy) and ULE (Unhealthy Life Expectancy).

We will then make assumptions to divide the retiree’s living spendings into basic and discretionary expenses so that our goal is to maximize the discretionary expenses while healthy, while keeping the retirement fund value above the basic expenses. The change of spending patterns depends on Markov simulations of HLE and ULE.

Annuities can provide risk-free income which is important to help eliminating ruin during retirement. We used payout ratio of SPIA to calculate how much basic expenses the retiree needs, and then purchase DIA to cover the bump up of retiree’s basic expenses while unhealthy.
We modeled the fund values using popular Geometric Brownian Motion models. Then we derived the solution for maximum annual withdrawals/discretionary expenses for each simulated scenario, considering both health and fund value simulations. The solution solves for maximum discretionary spending if the retiree starts healthy; solves for maximum withdrawal if starts unhealthy. We then choose the quantile (depending on the ruin probability) from the set of all solutions. Using numerical optimization methods, we will also obtain the optimal allocation of multiple assets in order to maximize the quantile solution.

Our research goes one step further from Xu 2018 [10] and proposes a recalibration technique that adjusts optimal spending annually based on historical actual investment returns and changes in health status.

All recalibration models can significantly reduce or even eliminate chance of ruin. But on the other hand, in order to maximize the annual withdrawals after recalibration, the process should have a sound mathematical basis, be consistent, be automated, and provide realistic solutions that a retiree can accept and live with. The Kalman Filter recalibration technique we are proposing satisfies all of these requirements.

Finally we will test our model performances compared to some traditional models by doing case studies under different market conditions. Moreover, we will show how we can use the model to provide bequest needs for a retiree and the projected liability cashflow for fund managers.
Chapter 2

Literature Review

Retirement financial planning is critical to a retiree’s life well-being. One of the most robust areas of research in the field of retirement financial planning is to determine how much a retiree can withdraw from their retirement and financial accounts. The research models are usually designed to optimize lifestyle under certain restrictions of ruin considering portfolio volatility, mortality and longevity.

Most of the emphasis of retirement planning models has been on fixed withdrawal rates from a retirement portfolio over a fixed time horizon. These models are static in the withdrawal rates(amounts) and also the planning horizon (See old models referenced in Stout 2006 [14]).

Stein 1998 [12] pointed out that retirement should be divided into three phases where life expenses gradually decrease due to worsening of health conditions. Based on this, Bengen 2001 [17] developed a dynamic model to compute withdrawals by randomly gradually decreasing withdrawals (if inflation not considered) and another dynamic model where withdrawals are randomly increased in bull market and are
decreased in bear market with an upper and lower bound for the change.

Pye 2001 [9] employs a withdrawal management technique that permits the retiree to withdraw the lower of the previous amount (in real terms) or the amortized current portfolio value using the plan length and expected portfolio rate of return. Assuming that security returns are distributed normally with a hypothetical 9% mean real rate of return and an 18% standard deviation, Pye reports that a 4.5% real withdrawal rate would require that 17% of retirees reduce real withdrawals within 20 years. Furthermore, the “worst off” 5% of retirees will need to reduce their withdrawals by at least 47%, even as the median portfolio value has approximately doubled.

Guyton 2004 [11] imposes systematic decision rules (restrictions) on withdrawal rate increases and subsequent make-ups for foregone withdrawal rate increases. For retirees willing to maintain the previous withdrawal rate in years after investment losses, and also cap inflationary increases in the withdrawal rate at 6% (without subsequent make-ups for either), the 40-year always safe initial withdrawal rate for an 80% equity portfolio is increased to 6.2%.

Stout 2006 [14] used periodic adjustments of retirement withdrawal rates based on both portfolio performance and remaining life expectancy, and Monte Carlo simulation of both investment returns and mortality. The inclusion of mortality in fixed planning horizon models reduces the probability of retirement-portfolio ruin by almost 50%. When compared to fixed withdrawal rate models, dynamic withdrawal management (Similar to Pye 2001 [9] with three types of controls on withdrawal rate changes) incorporating mortality reduces the probability of ruin by another 35–40% while increasing average lifetime withdrawal rates by nearly 50%.

Frank 2011 [8] developed three-dimensional distribution model that recognizes the transition from early retirement into later retirement. It used adjustment rules that
depend on how much the rate of return deviates from the historical averages.

Blanchett 2014 [6] showed that although the retiree consumption basket is likely to increase at a rate that is faster than general inflation – a fact that can largely be attributed to the higher weight to medical expenses for retirees – actual retiree spending tends to decline in retirement in real terms.

Suarez 2015 [16] develop iterative withdrawal strategies from retirement portfolios called the Perfect Withdrawal Amount, which can be considerably higher by avoiding the usual behavioural biases. It also followed a rule driven strategy by constructing and sequentially applying a probability distribution.

Xu 2018 [10] derived a quantile optimization method to solve for maximum withdrawals. It modeled a retiree’s health statuses as a Markov chain. It first derived a closed-form solution for each simulated scenario, and then chose the lower quantile value (according to the set ruin threshold) from solutions of all simulated scenarios as the final optimal solution. This method significantly reduced the computation time to be within seconds in R programming, compared to days or even weeks in the old trial-and-error method.

Kalman Filter is an iterative algorithm that is proven to be useful in solving continuous estimation problem like motion planning and control of vehicles and robotic, and also time series analysis, smoothing and forecasting. When applying to time series, Kalman Filter is typically applied to the actual values of time series instead of the underlying distribution parameters, as mentioned in Haravey 1989 [4] Chapter 4 and Chatfield 2003 [5]. In other words, the state variables of Kalman Filters are actual values of the time series. The common way to apply Kalman Filter to time series analysis can also be found in the document of the R package Kalman Filtering [3], where it combines Kalman Filter with ARIMA models.
Chapter 3

Methodology

3.1 Notations

3.1.1 Health conditions and Annual Expenses

i) \( B_h \): Basic expenses while healthy.

ii) \( B_u \): Basic expenses while unhealthy.

iii) \( E B_u \): Extra basic expenses while unhealthy. \( E B_u = B_u - B_h \).

iv) \( D_h \): Discretionary expenses while healthy.

v) \( D_u \): Discretionary expenses while unhealthy.

vi) \( T E_h \): Total expenses while healthy. \( T E_h = B_h + D_h \).

vii) \( T E_u \): Total expenses while unhealthy. \( T E_u = B_u + D_u \).
3.1.2 Assets and Annuities

i) \( m \): Number of different assets. Vectors in the following are all of length \( m \).

ii) \( T \): Initial total fund value of \( m \) assets.

iii) \( \vec{S}_i \): Asset values at year \( i \) before cash inflows and withdrawals.

iv) \( \vec{S}'_i \): Asset values at year \( i \) after cash inflows and withdrawals.

v) \( T_i \): Total fund value at year \( i \) before cash inflows and withdrawals.

vi) \( T'_i \): Total fund value at year \( i \) after cash inflows and withdrawals.

vii) \( || \cdot || \): \( L^1 \) norm of vectors. \( T_i = ||\vec{S}_i||, T'_i = ||\vec{S}'_i||. \)

viii) \( \vec{w} \): Allocation weights of assets. \( \vec{S}_i = T_i \vec{w}, \vec{S}'_i = T'_i \vec{w}. \)

ix) \( \vec{\mu} \): ‘Percentage drift’ in Geometric Brownian Motion modeling of assets.

x) \( \vec{\sigma} \): ‘Percentage volatility’ in Geometric Brownian Motion modeling of assets.

xi) \( \vec{a}, \vec{b} \): Constant vectors, \( \vec{a} = e^{\vec{\mu} - \frac{1}{2} \vec{\sigma}^2}, \vec{b} = e^{\vec{\sigma}}. \)

xii) \( \vec{Z}_i \): Standard Normal variables. \( \vec{S}_i = \vec{S}'_{i-1} \vec{a} \vec{b} \vec{Z}_i \), following GBM.

xiii) \( I_i \): Inflation rate at year \( i \).

xiv) \( A_d \): Purchasing amount of DIA at the initial year.

xv) \( p_d \): Annual payout ratio of DIA.

xvi) \( p_s \): Annual payout ratio of SPIA.

xvii) \( C_i \): Incoming cash flow at year \( i \), including social security income etc.
3.1.3 Kalman Filter

i) $x_k$: State vector at year $k$. Set to be $(\mu_k, \sigma_k)$.

ii) $\hat{x}_{k|k}$: Estimation of $x_k$ at year $k$. Also denoted as $\hat{x}_k$.

iii) $P_{k|k}$: Covariance matrix of $\hat{x}_k$. Also denoted as $P_k$.

iv) $z_k$: Measurement at year $k$.

v) $F$: State transition matrix.

vi) $Q$: Covariance matrix of state transition noise.


viii) $R_k$: Covariance matrix of measurement noise at year $k$.

3.2 Health Modeling and Spending patterns

3.2.1 Healthy and Unhealthy Life Expectancy

In order for our model to reflect the changing spending patterns while healthy and unhealthy, we need to differentiate between healthy and unhealthy statuses and get a good estimate of Healthy Life Expectancy (HLE) and Unhealthy Life Expectancy (ULE). (We will call the realization in each simulation Realized Life Expectancy(LR), Realized Healthy/Unhealthy Life Expectancy(HLR/ULR)) A natural idea is to use a non-homogeneous finite-state Markov chain with three states representing a retiree’s health status: healthy, unhealthy and dead. We make some assumption of the Markov model.
Assumption 3.2.1. Dead state can not transition to other two states and the simulation ends there. Unhealthy state can not transition to healthy state.

Here we used the same model as in Xu J. 2018 [10]. The transition probabilities can be inferred from mortality and morbility tables. The probabilities of healthy transitioning to dead depends on the retiree’s age, gender and being smoker or non-smoker, while probabilities from unhealthy to dead and from healthy to unhealthy depend on the retiree’s age and gender.

We can use this model to simulate the retiree’s future health statuses and have a reasonable model for retiree’s health status and thus a reliable prediction of HLE and ULE.

3.2.2 Basic Expenses and Discretionary Expenses

According to research and retirees’ need, the expenses of a retiree while healthy should be larger than the expenses while unhealthy so that the retiree can fully enjoy their physically vigorous years. In order to better model this, we can split the expenses into two parts – basic expenses and discretionary expenses – and make the following assumptions.

Assumption 3.2.2. A retiree has two kinds of life expenses: basic expenses and discretionary expenses. The spending pattern changes with respect to different health statuses.

(i) Basic expense while unhealthy is higher than that while healthy: $B_u > B_h$.

(ii) Discretionary expense while unhealthy is considered to be 0: $D_h > D_u = 0$.

(iii) Total expense while healthy is larger than that while unhealthy: $TE_h > TE_u$. 
Our model will give suggested amounts for $B_h$ and $B_u$ and optimize $D_h$. In case the retiree is already unhealthy or don’t want to use this pattern, we will assume constant future withdrawals and try to maximize the withdrawals.

### 3.2.3 Annuities: SPIA and DIA

SPIA (Single Premium Immediate Annuity) and DIA (Deferred Income Annuity) are two types of annuities. If one buys $1000$ SPIA and the payout ratio $p_s = 7.5\%$, then the benefit is immediately effective as $\$75$ each year until the death of the individual. If one buys $1000$ DIA and the payout ratio $p_d = 12\%$, then the benefit will be effective after a fixed length of time (usually set to be HLE) as $\$120$ each year until the death of the individual.

Reasonable basic expenses should depend on the retiree’s total asset value and health status. Because the retiree can not set this too high and should not set it too low, it would be a desired feature for our model to provide a reasonable suggested basic expenses $B_h$.

SPIA can provide the risk-free income if the retiree spend all the money buying SPIA. Similarly to that people typically set their basic expenses as a fixed percentage of salary, we can set the basic expenses as a fixed percentage of the payout from SPIA. For example, we can set the basic expenses while healthy to be $\frac{1}{3}$ of the payout money and set the basic expenses while unhealthy to be twice the amount of basic expenses while healthy, i.e.

\[
\frac{1}{2} B_u = EB_u = B_h = \frac{1}{3} T \cdot p_s,
\]  

(3.2.1)
where $T$ is the initial total fund value and $p_s$ is the payout ratio of SPIA. The
basic expenses calculated this way are expected to be reasonably small to avoid ruin
happening when without any discretionary expenses. If a financial advisor can find
other support for these factors or if there are specific needs from the retirees, they
can easily modify these values.

We can also buy annuities to cover some of the living expenses. Adding annuities
here will make our investment part able to take more risks and provide retiree basic
expenses in a less risky way, which is extremely important in the unlikely event when
investment market went really bad. However, we do not want to purchase too much
annuities. Because annuities has no liquidity and will leave retirees no money if they
die too early. Therefore our final solution is to purchase the exact amount of DIA so
that the bump up of basic expenses while unhealthy $EB_u$ is fully covered, i.e.

$$A_d \cdot p_d = EB_u, \text{ or } A_d = EB_u/p_d.$$  \hspace{1cm} (3.2.2)

\section*{3.3 Optimal Solution}

\subsection*{3.3.1 Asset Returns Modeling}

In order to perform annual recalibration and provide a more reliable suggested
withdrawal amount to the retiree, we need to model the asset returns better by
taking into account both their historical and current performance. Some research used
Normal Distribution to model the return ratio of assets, while Geometric Brownian
Motion (GBM) is more popular in modeling asset values, like stock prices in the
Black–Scholes model explained in Shreve book [15]. GBM treats the asset return
ratios as log normal and thus the asset values will always stay positive. Moreover, GBM is solution to the following stochastic differential equation which are used to model asset returns:

\[ dS_t = \mu S_t dt + \sigma S_t dW_t, \]

where \( W_t \) is a Standard Brownian Motion. By solving this equation we can get the following important property of GBM:

\[ S_T = S_t \exp \left\{ (\mu - \frac{1}{2} \sigma^2)(T - t) + \sigma \sqrt{T - t} Z \right\}, \]

where \( Z \sim N(0, 1) \) follows Standard Normal Distribution. The proof can be found in Shreve book [15].

When we have multiple assets and \( T - t = 1 \) year, we can use vector notations and denote \( \vec{a} = e^{\vec{\mu} - \frac{1}{2} \vec{\sigma}^2} \), \( \vec{b} = e^{\vec{\sigma}} \). We can then simplify the above formula as

\[ \vec{S}_{t+1} = \vec{S}_t \vec{a} \vec{b}^Z. \]

We can make the following assumptions about the changing patterns of our fund value.

**Assumption 3.3.1.** All cash inflows and withdrawals happen at the beginning of each year. Cash inflows will be used to invest at the beginning of that year. During each year, assets values follow Geometric Brownian Motion of different parameters with unknown dependencies.

Because of the spending patterns in our model, we can naturally define ruin as follows using basic expenses.
Definition 3.3.2. Ruin happens if at the beginning of any year before death, the total fund value $T_i$ after adding incoming cash goes below the basic expenses, which equals $B_h$ while healthy and $B_u$ while unhealthy.

Note that the above definition of ruin doesn’t take into account the withdrawals. However, since withdrawals are always larger than or equal to basic expenses, ruin happens if and only if it is not possible to withdraw from fund an equal or larger amount than basic expenses, i.e. the fund value at the beginning of year $i$ after cash inflow and withdrawal drop below zero. Thus, we have the following equivalent definition of ruin.

Definition 3.3.3. Ruin happens if $T_i' < 0$ for some year $i$ before death.

3.3.2 Quantile Optimization with Closed-Form Solution

For simplicity of formulas, we introduce two indicator variables here

i) $H_i$: $H_i = 1$ if retiree is healthy at year $i$; $H_i = 0$ if unhealthy at year $i$.

ii) $E_i$: Let $Y_{DIA}$ be the year when DIA first pays (usually HLE). $E_i = 1$ if $i \geq Y_{DIA}$; $E_i = 0$ if $i < Y_{DIA}$.

We will now denote $D_h$ as $D$ for simplicity and consistency of two different solutions (shown later). Then total expenses at year $i$ equals

$$B_h + EB_u + (D - EB_u)H_i.$$  

Suppose now we are looking at one simulated scenario $s$, which means that we know the assets returns (depending on $Z's$ in GBM simulation) and when the retiree
becomes unhealthy and dead. We will now solve for maximum discretionary spendings $D$ for this simulation so that ruin does not happen.

We initially have

$$T_0 = T - A_d$$

$$T'_0 = (T_0 + C_0 + A_{dpd}E_0 - B_h - EB_u - (D - EB_u)H_0)(1 + I_0)$$

$$\vec{S}'_0 = T'_0\vec{w}$$

$$\vec{S}'_1 = \vec{S}'_0\vec{a}\vec{b}\vec{Z}_i$$

$$T_1 = ||\vec{S}_1|| = ||T'_0\vec{a}\vec{b}\vec{Z}_i|| = T'_0||\vec{a}\vec{b}\vec{Z}_i||$$

Denote $F_i = ||\vec{a}\vec{b}\vec{Z}_i||$, then $T_1 = T'_0F_1$ and actually

$$T_i = T'_{i-1}F_i$$

We then have

$$T'_1 = [T'_0F_1 + C_1 + A_{dpd}E_1 - B_h - EB_u - (D - EB_u)H_1](1 + I_1)$$

$$= [(T_0 + C_0 + A_{dpd}E_0 - B_h - EB_u - (D - EB_u)H_0)(1 + I_0)F_1$$

$$+ C_1 + A_{dpd}E_1 - B_h - EB_u - (D - EB_u)H_1](1 + I_1)$$

$$= T_0F_1(1 + I_0)(1 + I_1)$$

$$+ C_0F_1(1 + I_0)(1 + I_1) + C_1(1 + I_1)$$

$$- (B_h + EB_u)[F_1(1 + I_0)(1 + I_1) + (1 + I_1)]$$

$$+ A_{dpd}[E_0F_1(1 + I_0)(1 + I_1) + E_1(1 + I_1)]$$

$$- (D - EB_u)[H_0F_1(1 + I_0)(1 + I_1) + H_1(1 + I_1)]$$
\[ T'_2 = [T'_1 F_2 + C_2 + A_d p_d E_2 - B_h - E B_u - (D - E B_u) H_2] (1 + I_2) \]
\[ = T_0 F_1 F_2 (1 + I_0)(1 + I_1)(1 + I_2) \]
\[ + C_0 F_1 F_2 (1 + I_0)(1 + I_1)(1 + I_2) + C_1 F_2 (1 + I_1)(1 + I_2) + C_2 (1 + I_2) \]
\[ - (B_h + E B_u) [F_1 F_2 (1 + I_0)(1 + I_1)(1 + I_2) + F_2 (1 + I_1)(1 + I_2) + (1 + I_2)] \]
\[ + A_d p_d [E_0 F_1 F_2 (1 + I_0)(1 + I_1)(1 + I_2) + E_1 F_2 (1 + I_1)(1 + I_2) + E_2 (1 + I_2)] \]
\[ - (D - E B_u) [H_0 F_1 F_2 (1 + I_0)(1 + I_1)(1 + I_2) + H_1 F_2 (1 + I_1)(1 + I_2) + H_2 (1 + I_2)] \]

Let \( G^i_j = (1 + I_k) \prod_{k>j} F_k (1 + I_{k-1}) \) and \( G^i_i = (1 + I_i) \), then

\[ T'_i = (T - A_d) G^i_0 \]
\[ + (C_0 G^i_0 + C_1 G^i_1 + \ldots + C_i G^i_i) \]
\[ - (B_h + E B_u) (G^i_0 + G^i_1 + \ldots + G^i_i) \]
\[ + A_d p_d (E_0 G^i_0 + E_1 G^i_1 + \ldots + E_i G^i_i) \]
\[ - (D - E B_u) (H_0 G^i_0 + H_1 G^i_1 + \ldots + H_i G^i_i) \]

According to Definition 3.3.3 of ruin, to maximize \( D \) and spend all our money at year \( i \), we need to set \( T'_i = 0 \) and solve for \( D \) in this simulated scenario \( s \). We can therefore get

\[ D^i_s = E B_u + \frac{1}{H_0 G^i_0 + H_1 G^i_1 + \ldots + H_i G^i_i} \left\{ (T - A_d) G^i_0 \\ + (C_0 G^i_0 + C_1 G^i_1 + \ldots + C_i G^i_i) \\ - (B_h + E B_u) (G^i_0 + G^i_1 + \ldots + G^i_i) \\ + A_d p_d (E_0 G^i_0 + E_1 G^i_1 + \ldots + E_i G^i_i) \right\} \] (3.3.1)
If the retiree becomes unhealthy or disabled at some year, then we will not distinguish his future health status between healthy and unhealthy, and thus will not assume the change in spending patterns in the future financial planning strategies. In this case, instead of assuming fixed basic expenses and trying to maximize the discretionary expenses while healthy, we will simply assume constant withdrawal for each simulated future year and try to maximize this withdrawal amount. The withdrawal should be larger than basic expenses. For simplicity of notation, we will also denote the amount of future withdrawals less basic expenses as \( D \), i.e. withdrawal \( = D + B_u = D + B_h + EB_u \). The derivation process is very similar to the one above but without the health condition indicator \( H'_i \). The closed-form solution now will be

\[
D^i_s = \frac{1}{G^i_0 + G^i_1 + \cdots + G^i_T} \{ (T - A_d)G^i_0 \\
+ (C_0 G^i_0 + C_1 G^i_1 + \cdots + C_i G^i_i) \\
+ A_d p_d (E_0 G^i_0 + E_1 G^i_1 + \cdots + E_i G^i_i) \}
\]

\[
- (B_h + EB_u)
\]

(3.3.2)

For each simulated scenario \( s \), we should let \( i \) equal to LR and solve for optimal withdraw, i.e.

\[
D_s = D_{sLR}.
\]

However, we need to worry about the case when our asset value falls below zero then get back to being positive. This only happens if at the beginning of one year withdrawal is greater than income and at some future year withdrawal is smaller than income. In this case, we need to make sure that ruin does not happen at any year
before death. We can solve for $D$ using the solutions by setting $i$ equal to $1, 2, \cdots, LR$ and use the smallest $D$ as the solution for this simulated scenario $s$.

$$D_s = \min(D_1^s, D_2^s, \cdots, D_{LR}^s)$$

Now we have a closed-form solution for each simulation $s$ with known future health conditions and assets returns. Suppose for example that our ruin threshold is set to be $5\%$. In order to maximize the withdrawals subject to the restriction that ruin will happen only in $5\%$ of all simulated scenarios, we need to simply compute the $5\%$ quantile in the set of solutions to all simulations.

$$D = \text{quantile}(5\%, \{D_s \mid \text{for all simulation } s\}) \quad (3.3.3)$$

After deriving the solution for optimal withdrawals, we can test it by setting arbitrary testing values for $\vec{w}, T$ and health conditions then going forward in time, simulate asset paths and take the withdrawal amount calculated from the solution. As shown in Figure 3.3.1, we can see that the fund value will become 0 at death time.

### 3.3.3 Multiple Assets: Allocation and Covariance

To achieve best fund performance with lower risk, people will create a portfolio of more than one asset, which is already the case in all the variables and formulas above. The solution given above in (3.3.1) and (3.3.2) used the asset allocation vector $\vec{w}$ (contained in $G^i_j$'s). In order to maximize the solution $D$ in (3.3.3), we need to determine the specific allocation weights in $\vec{w}$.

Optimal assets allocation is an important topic studied for a long time in Financial
mathematics. The famous theory of efficient frontier can be found in many books, for example Elton 2011 [7]. However, the optimal allocation is usually in the sense of maximizing expectation under constraints of variance. Our problem is about the optimization in a random time horizon with a set ruin probability where these kind of theory does not work.

We can however easily formulate the problem as a typical constrained optimization problem with **objective function** given in Equation (3.3.3):

\[ D = \text{quantile}(5\%, \{D_p\text{ for all simulation }p\}) \]

with **input variables** being \( \bar{w} \) under the **constraints** \( \bar{w} \geq 0 \) and \( ||\bar{w}|| = 1 \).

Now that we can resort to plenty of numerical methods, like generalized Lagrange Multipliers method and gradient based optimization, provided in many numerical packages, like “Rsolnp” in R and “scipy.optimize.minimize” in Python. When precision permitted and number of assets is small enough, we can use the simpliest
brute-force grid search method to search for the best assets allocation to make the searching even faster and to obtain global optimal solution.

Moreover, since we have multiple assets, it makes sense to assume dependencies among different assets. The covariance matrix can be obtained directly from financial institutions and companies, or computed based on historical data. We then used the popular method based on the Cholesky decomposition of the covariance matrix (Scipy documentation [1]) to generate correlated random samples (the Z’s in GBM).

3.4 Kalman Filter

As mentioned in Chapter 2 Literature Review, recalibration techniques proposed earlier (Bengen 2001 [17], Pye 2001 [9], Guyton 2004 [11], Stout 2006 [14], Suarez 2015 [16], Frank 2011 [8]) are somewhat naive and random. We will apply Kalman Filter for the recalibration process to make it more mathematically rigorous.

3.4.1 Derivation and Properties

Kalman Filter is a method that has been shown to perform well in solving the so-called Estimation Problem in time series analysis and communication engineering. This is to seek the continual estimation of a set of parameters (signal of interest x(t)) whose values change over time, by combining a set of observations or measurements z(t) which contain information about x(t).

An estimator is considered good if it is unbiased and can minimize some loss function of estimation.
Definition 3.4.1 (Unbiased Condition).

\[ \mathbb{E}[\hat{x}] = x \]

A function \( L \) is a general loss function if it satisfies the following definition:

Definition 3.4.2 (Loss Function).

\[ L(\epsilon_2) \geq L(\epsilon_1) \geq 0 \text{ when } \epsilon_2 \geq \epsilon_1 \geq 0 \]

and \( L(\epsilon) = L(-\epsilon) \).

We need the following definitions in order to derive the algorithm of Kalman Filter.

Definition 3.4.3 (Conditional Distribution Function).

\[ F(\xi) = \mathbb{P}[x(t_1) \leq \xi | z(t_0) = \eta(t_0), \cdots , z(t) = \eta(t)] \]

Definition 3.4.4 (Average Loss).

\[ \mathbb{E}\{L[\hat{x}(t_1) - x(t_1)]\} = \mathbb{E}[\mathbb{E}\{L[\hat{x}(t_1) - x(t_1)]|z(t_0), \cdots , z(t)\}] \]

Kalman 1960 [13] showed the following theorems.

Theorem 3.4.5. If the conditional distribution function \( F \) satisfies:

\[ F(\xi - \bar{\xi}) = 1 - F(\bar{\xi} - \xi) \]

\[ F(\lambda \xi_1 + (1 - \lambda)\xi_2) \leq F(\xi_1) + (1 - \lambda)F(\xi_2) \]
for all $\xi_1, \xi_2 \leq \bar{\xi}$ and $0 \leq \lambda \leq 1$, then the random variable $\hat{x}^*$ which minimizes the average loss $\mathbb{E}\{L[\hat{x}(t_1) - x(t_1)]\}$ is the conditional expectation

$$\hat{x}^*(t_1|t) = \mathbb{E}[x(t_1)|z(t_0), \cdots, z(t)].$$

More useful results from Kalman 1960 [13] are the following:

**Corollary 3.4.6.** If the random processes $\{x(t)\}, \{w(t)\}$ and $\{z(t)\}$ are Gaussian, the previous theorem holds.

**Theorem 3.4.7.** If $L(\epsilon) = \epsilon^2$, then the previous theorem is true without the symmetry and convexity assumptions.

Following is a well-known result about linearity of estimators:

**Theorem 3.4.8.** If the random processes $\{x(t)\}, \{z(t)\}$ are Gaussian, the optimal estimator is a linear estimator.

Now before deriving the Kalman Filter algorithm, we need to rephrase the estimation more specifically. If assuming we can model our problem as a two-step iterative process. The **state transition equation** is given by

$$x_{k+1} = F_k x_k + G_k u_k + w_k,$$  \hspace{1cm} (3.4.1)

where $x_k$ is the state at time $k$, $u_k$ is an input control vector, $w_k$ is additive system or process noise, $G_k$ is the input transition matrix and $F_k$ is the state transition matrix.

And the **measurement equation** is given by

$$z_k = H_k x_k + v_k,$$  \hspace{1cm} (3.4.2)
where $z_k$ is the observation or measurement made at time $k$, $x_k$ is the state at time $k$, $H_k$ is the measurement matrix and $v_k$ is additive measurement noise.

We will need to make assumptions of distribution of variables.

**Assumption 3.4.9.** The process and measurement noise random processes $w_k$ and $v_k$ are uncorrelated, zero-mean white noise with known covariance matrices, $Q_k$ and $R_k$.

The initial system state, $x_0$ is a random vector that is uncorrelated to both the system and measurement noise processes, and has a known mean and covariance matrix $\hat{x}_{0|0}$ and $P_{0|0}$.

We will now present the derivation of Kalman Filter here as a Linear Minimum Variance of Error (LMV) filter (i.e. it is the optimal linear filter over the class of all linear filters) and an unbiased filter.

In the case of the state vector $x$ and the observations $z$ are jointly Gaussian distributed, the Minimum Variance of Error Estimator (MVE) is a linear function of $z$ and thus is also a LMV.

Given a set of observations $z_1, \cdots, z_{k+1}$, now our task is to determine estimation $\hat{x}_{k+1}$ as a linear function of $z$’s that minimises the expectation of the squared-error loss function

$$E[||x_{k+1} - \hat{x}_{k+1}||^2] = E[(x_{k+1} - \hat{x}_{k+1})^T(x_{k+1} - \hat{x}_{k+1})]$$

In other words, we will prove the following theorem.

**Theorem 3.4.10.** Under Assumption 3.4.9, the following Kalman Filter algorithm provides an unbiased minimum variance of error estimator where the model satisfies
Equation 3.4.1 and 3.4.2.

Given the estimation \( \hat{x}_{k|k} \) and its covariance matrix \( P_{k|k} \) at time step \( k \), following is the Kalman Filter algorithm at time step \( k+1 \) based on new measurement \( z_{k+1} \), which gives the estimation \( \hat{x}_{k+1|k+1} \) and its covariance matrix \( P_{k+1|k+1} \).

**State Prediction**

\[
\hat{x}_{k+1|k} = F_k \hat{x}_{k|k} + G_k u_k \\

P_{k+1|k} = F_k P_{k|k} F_k^T + Q_k
\]

**Measurement Update**

\[
K_{k+1} = P_{k+1|k} H_{k+1}^T [H_k P_{k+1|k} H_k^T + R_{k+1}]^{-1} \\

\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1} [z_{k+1} - H_{k+1} \hat{x}_{k+1|k}] \\

P_{k+1|k+1} = (I - K_{k+1} H_{k+1}) P_{k+1|k} (I - K_{k+1} H_{k+1})^T + K_{k+1} R_{k+1} K_{k+1}^T
\]

**Proof.** (Derivation of Kalman Filter)

Denote \( \hat{x}_{k|i} \) as the estimation of \( x \) at time \( k \) based on time \( i \), \( k \geq i \).

The so-called one-step-ahead prediction or simply a prediction

\[
\hat{x}_{k+1|k} = \mathbb{E} [x_{k+1} | z_1, \cdots, z_k] = \mathbb{E} [x_{k+1} | Z^k] = \mathbb{E} [x_{k+1} | F_k x_k + G_k u_k + w_k | Z^k] = F_k \hat{x}_{k|k} + G_k u_k
\]
The estimate variance

\[
P_{k+1|k} = \mathbb{E}[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T | Z^k] = \mathbb{E}[(F_k x_k - F_k \hat{x}_{k|k} + w_k)(F_k x_k - F_k \hat{x}_{k|k} + w_k)^T | Z^k] = F_k \mathbb{E}[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T | Z^k] F_k^T + \mathbb{E}[w_k w_k^T] = F_k P_{k|k} F_k^T + Q_k
\]

Suppose now we have the observation \( z_{k+1} \). Since the estimate is a linear weighted sum of the prediction and the new observation and can be described by the equation

\[
\hat{x}_{k+1|k+1} = K'_{k+1} \hat{x}_{k+1|k} + K_{k+1} z_{k+1}
\]

Our problem now is reduced to finding the \( K'_{k+1} \) and \( K_{k+1} \) that minimise the conditional mean squared estimation error.

Now we will require the unbiased uncondition on the estimators.

Argue by induction and assume that \( x_{k|k} \) is an unbiased estimate. Then

\[
\mathbb{E}[\hat{x}_{k+1|k+1}] = \mathbb{E}[K'_{k+1} \hat{x}_{k+1|k} + K_{k+1} H_{k+1} x_{k+1} + K_{k+1} u_{k+1}] = K'_{k+1} \mathbb{E}[\hat{x}_{k+1|k}] + K_{k+1} H_{k+1} \mathbb{E}[x_{k+1}] + K_{k+1} \mathbb{E}[u_{k+1}] = K'_{k+1} \mathbb{E}[F_k \hat{x}_{k|k}] + G_{k} u_{k} + K_{k+1} H_{k+1} \mathbb{E}[x_{k+1}] = K'_{k+1} (F_k \mathbb{E}[\hat{x}_{k|k}] + G_{k} u_{k}) + K_{k+1} H_{k+1} \mathbb{E}[x_{k+1}] = K'_{k+1} \mathbb{E}[x_{k+1}] + K_{k+1} H_{k+1} \mathbb{E}[x_{k+1}]
\]

Thus, in order for \( \hat{x}_{k+1|k+1} \) to be unbiased, we need

\[
K'_{k+1} + K_{k+1} H_{k+1} = I
\]
Therefore, by requiring our estimate to be unbiased, we have:

\[
\hat{x}_{k+1|k+1} = (I - K_{k+1}H_{k+1})\hat{x}_{k+1|k} + K_{k+1}z_{k+1}
\]

\[
= \hat{x}_{k+1|k} + K_{k+1}[z_{k+1} - H_{k+1}\hat{x}_{k+1|k}]
\]

where \( K \) is known as the Kalman gain and is to be determined.

Denote \( \tilde{x}_{k+1|k+1} \) to be the estimation error \( x_{k+1} - \hat{x}_{k+1|k+1} \).

\[
P_{k+1|k+1} = \mathbb{E}[(x_{k+1} - \tilde{x}_{k+1|k+1})(x_{k+1} - \tilde{x}_{k+1|k+1})^T]
\]

\[
= (I - K_{k+1}H_{k+1})\mathbb{E}[\tilde{x}_{k+1|k}\tilde{x}_{k+1|k}^T](I - K_{k+1}H_{k+1})^T
\]

\[+ K_{k+1}\mathbb{E}[v_{k+1}v_{k+1}^T]K_{k+1}^T
\]

\[+ 2(I - K_{k+1}H_{k+1})\mathbb{E}[\tilde{x}_{k+1|k}v_{k+1}^T]K_{k+1}^T
\]

\[
= (I - K_{k+1}H_{k+1})P_{k+1|k}(I - K_{k+1}H_{k+1})^T + K_{k+1}R_{k+1}K_{k+1}^T
\]

The Kalman Gain

\[
K_{k+1} = \text{argmin} \ \mathbb{E}[\tilde{x}_{k+1|k+1}^T\tilde{x}_{k+1|k+1}|Z^{k+1}]
\]

\[= \text{argmin} \ \text{trace}(\mathbb{E}[\tilde{x}_{k+1|k+1}\tilde{x}_{k+1|k+1}^T|Z^{k+1}])
\]

\[= \text{argmin} \ \text{trace}(P_{k+1|k+1})
\]

In order to minimize the variance, we differentiate with respect to \( K_{k+1} \) and set equal to zero:

\[
-2(I - K_{k+1}H_{k+1})P_{k+1|k}H_{k+1}^T + 2K_{k+1}R_{k+1} = 0
\]

\[
K_{k+1} = P_{k+1|k}H_{k+1}^T[H_{k+1}P_{k+1|k}H_{k+1}^T + R_{k+1}]^{-1}
\]
In summary, the Kalman Filter is an iterative process to estimate some state variables (denoted \(x\)), with new measurements (denoted \(z\)) coming in at each time step. Both the states and measurement contain uncertainties and noise.

Prediction step: Use \(x_{t-1}\) to predict \(x_t\) along with its uncertainties.

Update step: Use the predicted \(x_t\) and new measurement \(z_t\) to update \(x_t\) and its uncertainties. The update uses a weighted average, with more weight being given to estimates with higher certainty.

### 3.4.2 Kalman Filter Set Up

When we use GBM to model the asset values, we will need better estimation of \(\mu\) and \(\sigma\) so that better future simulations can be used to compute a more realistic solution of optimal withdrawal or discretionary expenses. To make our model dynamic, we can continuously estimate these parameters using Kalman Filter. Future simulations of returns are then based on these new estimation of parameters at each recalibration step. We are now going to set up the Kalman Filter.

We need to firstly choose the time interval length of our measurement update. Here we choose to recalibrate annually.

We then need to define state space and measurements, together with the state transition function and measurement update function, as in Equation 3.4.1 and 3.4.2.
State vector

\[ x_k = (\mu_k, \sigma_k^2) \]

and the estimation \( \hat{x}_k \) at year \( k \) is 2-dimensional multivariate Gaussian distributed centered at \( x_k \) with covariance matrix denoted \( P_k \). Note that we used \((\mu_k, \sigma_k^2)\) instead of \((\mu_k, \sigma_k)\) in order to make the later measurement update equation linear. Otherwise, we have to use Extended Kalman Filter or Unscented Kalman Filter to get approximated results.

Assume that the GBM parameters stay unchanged for each subsequent year with possible variations modeled by time homogeneous white noise. Thus the state transition function is:

\[ x_{k+1|k} = F x_k + w, \quad \text{where} \quad F = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad w \sim N(0, Q). \]

We need new measurements \( z_k \) every year that are related to \( x_k \), since \( \mu \) and \( \sigma \) cannot be observed directly. Note that the log ratio of a GBM is Gaussian, or equivalently, the difference of log returns of GBM is Gaussian. Therefore, we can use the daily returns from last year to get estimation of the mean and variance of this Gaussian variable as \( z_k \) that is related to \( x_k \).

More specifically, to use daily returns as data points for estimation, let \( dt \) be one day, i.e. \( 1/(\text{number of trading days}) \) year. Since

\[ i_t = \log(S_t) - \log(S_{t-dt}) \sim N((\mu_k - \frac{1}{2}\sigma_k^2)dt, \sigma_k^2 dt), \]
we can define **measurements** as the mean and variance of $i_t$

$$z_k = \left((\mu_k - \frac{1}{2}\sigma_k^2)dt, \sigma_k^2 dt\right),$$

with **measurement function**

$$z_k = Hx_k + v_k, \text{ where } H = \begin{pmatrix} dt & -\frac{1}{2}dt \\ 0 & dt \end{pmatrix}, \text{ and } v_k \sim N(0, R_k).$$

Finally, we need to consider what values we should use for the hyperparameters in Kalman Filter, containing the initial values of estimation and its covariance matrix, and also the two covariance matrices of white noises.

i) $\hat{x}_0$ : Estimate of $(\mu_0, \sigma_0^2)$. We used historical data of assets to estimate it, by using the random variable $i_t$ and the same method as above.

ii) $P_0$ : Assume $\mu$ and $\sigma$ are independent. The larger the variances, the less sure we are about the initial estimate $\hat{x}_0$, and thus comparably more sure about the measurements followed. Need to be tuned.

iii) $Q$ : Variance of state transition noise. Imply the variability of $x$ as time goes by. The larger the values, the more sure we are about the new measurement $z_k$ compared to the predicted estimate $\hat{x}_{k|k-1}$. Need to be tuned.

iv) $R_k$ : Variance of measurement noise. Since our measurement comes from the estimation of mean and variance of the normal variable $i_t$, we can use sampling distribution of the estimation $\hat{i}_t$. In general, for a normal variable $N(\mu, \sigma^2)$, the sampling distribution of estimation of the estimation of $\mu$ and
σ are given by $\frac{X - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$ $\Rightarrow$ var($\bar{X}$) = $\frac{\sigma^2}{n}$ and $\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2 \Rightarrow$ var($S^2$) = $\frac{2\sigma^4}{n-1}$. However, since the value of $R_k$ is compared to $Q$ to indicate the importance of measurement v.s. prediction, we need to also tune the value of $R_k$. We can do this by introducing a constant multiplier $C$, so that $R_k = \begin{pmatrix} C \frac{\sigma^2}{n} & 0 \\ 0 & C \frac{2\sigma^4}{n-1} \end{pmatrix}$, where $\sigma^2$ here is the variance of $i_t$.

In order to tune these parameters, we will look at the performance of estimations of $x_k = (\mu_k, \sigma_k)$. We simulated a lot of paths for asset returns with different true $\mu$ and $\sigma$ values. After randomly given the initial estimate $\hat{x}_0$, we tuned the parameters to make the estimated values close enough to the true value and not oscillating too much, but also be able to reflect the recent market performances. After a lot of tuning, we arrived at a certain set of parameters that can make the estimation perform well on all simulated paths with different true $\mu$ and $\sigma$ values. Figure 3.4.1 is an interesting example of one simulated path of 30-year asset returns, generated by the true $\mu = 0.03$ and $\sigma^2 = 0.09$. Figure 3.4.2 showed the estimation performance where the estimated values are represented by dotted blue lines and true values by black horizontal lines.
To model more than one assets, our approach is to create one Kalman Filter for each asset. It is also possible to set up just one Kalman Filter for all assets with $2 \times m$ state variables where $m$ is the number of assets. But we found it very hard for parameter tuning and the performance is unpredictable when $m$ is large and we have large vectors/matrices. However, the good parameters we got above suits all reasonable $\mu$ and $\sigma$ values and thus can be copied directly to all assets to make life much easier. Although the Kalman Filter does not capture dependencies among assets in its set up, any dependencies are actually reflected in the updated results after the independent estimation processes.
Chapter 4

Model Results

4.1 Model Summary

We can now summarize the assumptions and set-ups of our model

(i) Assets follow Geometric Brownian Motion with dependencies.

(ii) Health condition is modeled by the Markov Chain.

(iii) Spending patterns are explained in Section 3.2.2 and Section 3.2.3.

(iv) The algorithm will find optimal assets allocation in order to maximize $D$ in the sense of quantile optimization using closed-form solutions showed in Section 3.3.

Inputs to the model are:

(i) Initial fund value $T$.

(ii) Health inputs: gender, age, smoker/nonsmoker.
(iii) Initial values for Kalman Filter: $\mu_0$, $\sigma_0$, $P_0$.

(iv) Payout ratios of annuities $p_s, p_d$; inflation rates $I_i$; incomes $C_i$.

(v) Need daily asset values from last year to perform Kalman Filter recalibration.

**Outputs** from the model are:

(i) $D$. Provides a new $D$ every year if annually recalibrated.

(ii) Basic expenses $B_h, B_u$.

(iii) Assets allocation $\vec{w}$. Provides a new $\vec{w}$ every year if annually recalibrated.

(iv) $A_d$: amount of DIA to buy at the start.

So here is how a financial advisor should use our model.

At the initial year, the agent collects information from the retiree, including initial fund value, gender, age, smoking/non-smoking. The agent needs to input historical data of the assets in the fund so that the model can get initial estimates of Kalman Filter parameters $\mu_0, \sigma_0, P_0$. The agent can also input the values of $\mu_0, \sigma_0, P_0$ if already has a confident estimate. The agent also needs to input the payout ratio of SPIA and DIA, inflation rate $I_i$ and also future incoming cash flows $C_i$. Then the model will provide the suggested amount of basic expenses while healthy $B_h$ and basic expenses while unhealthy $B_u$, how much DIA to buy in the first year $A_d$, asset allocation $\vec{w}$ and also suggested discretionary expenses $D$.

After that, every year the agent will provide the model with daily values of the assets in the fund from previous year so that the model can perform recalibration and output a new $D$ and $\vec{w}$. The agent can also change certain inputs if needed, including
the retiree being smoking or non-smoking, future incoming cash flows $C_i$ and inflation rates $I_i$.

If at some year, the retiree becomes unhealthy, the agent needs to tell the model so that it no longer assumes changing of spending patterns in the future and will output suggested constant annual withdrawal.

### 4.2 Simulation Method

We will now explain how simulations and annual recalibrations are done in order to obtain results shown in Section 4.3 for model performances comparison.

Suppose we have $m$ assets. For each simulated scenario, we will simulate $m$ many GBM paths representing return of assets, which depend on the value of $\vec{\mu}, \vec{\sigma}, T, \vec{w}$ and are generated by sequences of Standard Normal Variables $Z'$s considering their dependencies. We only care about the part before death, and whether the spending patterns are that when healthy or unhealthy.

\[
\text{GBM paths } + (\text{HLR, ULR}) \xrightarrow{\text{Solution (3.3.3)}} D
\]

Given $\vec{\mu}, \vec{\sigma}$ and $T$, suppose ruin threshold is set to be 5%, we simulate $1,000 \times m$ GBM paths and 1,000 different health conditions, i.e. (HLR, ULR)'s. Then we can have $1,000 \times 1,000$ $D$'s as solutions to each simulated scenario. We will search for an optimal assets allocation $\vec{w}$ so that the 5% quantile $D$ is maximized.

To test model performances, especially for model with recalibrations, we need to have enough simulated scenarios that are considered “True”. Model applied at year $i$ can and can only use the information of “True” simulations up to year $i$. The tricky
part is that when we solve for $D$, we need simulations after current year $i$. These future simulations are “Fake” and are only used to solve for $D$, and thus should be a different set of simulations from the existing “True” simulations.

Suppose that the true $\vec{\mu}$ and $\vec{\sigma}$ are fixed and we generate one simulated “true” scenario with $m$ GBM paths and one (HLR, ULR) pair, following is a simple illustration of what happens in each year in this scenario.

Year 1: initial estimate $\hat{\mu}_1, \hat{\sigma}_1$; output $D$ for year 1; simulate daily returns $DR_1$.

Year 2: new estimate $\hat{\mu}_2, \hat{\sigma}_2$ based on $DR_1$; output $D$ for year 2; simulate daily returns $DR_2$.

Year 3: new estimate $\hat{\mu}_3, \hat{\sigma}_3$ based on $DR_2$; output $D$ for year 3; simulate daily returns $DR_3$.

\[ \cdots \]

Figure 4.2.1 shows these actions in more details for one simulated scenario in year $i$. Inside the dotted circle is what happened before year $i$ when we generate GBM paths in year $i - 1$. Outside the dotted circle is what happened at the beginning of year $i$ when we use Kalman Filter recalibration to get new estimated $\vec{\mu}$ and $\vec{\sigma}$ then solve for the optimal solution of $D$ and allocation weight $\vec{w}$.

To test for model performance, for example, we can then simulate a total of $1,000 \times 1,000$ “True” simulations and perform these actions $1,000 \times 1,000$ times.

### 4.3 Case Studies

We will compare our model (in Section 4.3.5) to some traditional and newer models by performing case studies on variant inputs and evaluate the model performances.
Following are inputs to the models in all case studies.

(i) Initial asset value $T =$1,000,000.

(ii) Retiree is female, age 65, non-smoker

(iii) We will consider 3 assets, from safe to moderate then risky. We chose from historical data as reasonable values. $\bar{\mu} = (0.02127213, 0.05145323, 0.06372564)$, $\bar{\sigma} = (5.8291644e-05, 1.0429876e-02, 2.5947163e-02)$.

(iv) $p_s = 6.38\%$ and $p_d = 25.5\%$ quoted from ImmediateAnnuities.com [2]. Using Equation (3.2.1) and (3.2.2), we know $B_h = EB_u = $21,267 and we should buy $A_d = $83,400 DIA which pays after 17 years.

(v) $I_i$ are set to be 0 for all $i$. Since we want to compare outputs from different models, we will need to use inflation-adjusted amounts to compare. Thus it is
better to not consider inflation at all when evaluating model performance.

(vi) We will only consider social security income as the only source of cash inflow in our case studies. Since the average social security income for year 2019 is $1,461 per month, which is $17,532 for 12 months, we will let $C$ equal to $17,532$.

The distribution of health simulations would be steady if we have 1,000 simulation, as shown in Figure 4.3.1.

The distribution of asset return simulations would be steady if we have 1,000 simulation, as shown in Figure 4.3.2.

Therefore we will have a total of $1,000 \times 1,000$ simulated "True" scenarios to test our model performances. Then each time when we solve for optimal discretionary spendings/withdrawals, we will need another set of $1,000 \times 1,000$ simulated future scenarios.

We will compute discretionary expenses or annual withdrawals using a set of simulations then evaluate the averaging performances of models when market is good,
moderate and bad, based on another set of simulations. We will do this by modifying the \( \mu \)'s when we generate the true paths of asset returns using GBM, i.e., adding 0.005 to generate good scenarios, minus 0.005 to generate bad scenarios and keep \( \mu \) unchanged for moderate scenarios.

4.3.1 Model 1: Fixed Death Time

This is the most commonly used traditional model when we assumed the retiree dies at a fixed age, age 95. And we do not consider changing spending patterns in different health status and we do not buy any DIA to cover basic expenses. We only need 1,000 simulations of fund values because death time is fixed.

We first calculate the quantile solution of withdrawals based on a set of fund values simulations. Then we test how much money would be left when the retiree is dead in each simulation when market performance is good, moderate and bad. Finally, we compute ruin by using the number of simulated scenarios when ruin happens dividing by the total number of simulations. Following table shows the withdrawals, average
ending fund value and the ruin probability.

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Good</th>
<th>Moderate</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual withdrawals</td>
<td>$82,815</td>
<td>$82,815</td>
<td>$82,815</td>
</tr>
<tr>
<td>Average ending fund value</td>
<td>$1,119,896</td>
<td>$671,162</td>
<td>$304,323</td>
</tr>
<tr>
<td>Ruin probability</td>
<td>0.82%</td>
<td>5.00%</td>
<td>21.46%</td>
</tr>
</tbody>
</table>

4.3.2 Model 2: Random Death Time

In this model, we will add randomness in the retiree's death time. We do this by using the probabilities from annuity tables to simulate death. Same as Model 1, we do not consider changing spending patterns in different health status and we do not buy any DIA to cover basic expenses.

We used $1000 \times 1000$ simulations to calculate the optimal solutions, then we test how this solution performs on other sets of simulations when market is good, moderate or bad. Following is the table showing withdrawals, average ending fund value and the ruin probability.

<table>
<thead>
<tr>
<th>Model 2</th>
<th>Good</th>
<th>Moderate</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual withdrawals</td>
<td>$84,066</td>
<td>$84,066</td>
<td>$84,066</td>
</tr>
<tr>
<td>Average ending fund value</td>
<td>$1,018,218</td>
<td>$730,340</td>
<td>$491,336</td>
</tr>
<tr>
<td>Ruin probability</td>
<td>1.23%</td>
<td>5.00%</td>
<td>12.91%</td>
</tr>
</tbody>
</table>

4.3.3 Model 3: HLE and ULE

This is the model we talked about in Chapter 3 without annual recalibration. We will add HLE and ULE Markov chain simulations and consider the changing patterns
in different health statuses. We will model living expenses as basic plus discretionary and try to maximize the discretionary expenses while retiree is healthy. We will calculate basic expenses based on the payout from SPIA and buy DIA to cover the bump up of basic expenses while unhealthy $EB_u$.

We used $1000 \times 1000$ simulations to calculate the optimal solutions of discretionary expenses while healthy, then we test how this solution performs on other sets of simulations when market is good, moderate or bad. Following is the table showing withdrawals while healthy, average ending fund value and the ruin probability. Note that withdrawals while unhealthy are constant because it only contains basic expenses $B_u$ and is equal to $42534$, as explained in Section 4.3 (iv).

<table>
<thead>
<tr>
<th></th>
<th>Model 3</th>
<th>Good</th>
<th>Moderate</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Withdrawals while healthy</td>
<td>$86,180$</td>
<td>$86,180$</td>
<td>$86,180$</td>
<td></td>
</tr>
<tr>
<td>Average ending fund value</td>
<td>$1,022,397$</td>
<td>$772,747$</td>
<td>$564,859$</td>
<td></td>
</tr>
<tr>
<td>Ruin probability</td>
<td>1.44%</td>
<td>5.00%</td>
<td>11.90%</td>
<td></td>
</tr>
</tbody>
</table>

### 4.3.4 Model 4: Naive Recalibration

In this model, we will simply rerun Model 3 every year as a comparison to our final model, Model 5, which uses Kalman Filter to do annual recalibration. At the beginning of each year, the starting fund value and the age of retiree are changed as inputs to the optimal quantile solution. Also, we need to determine whether to use (3.3.1) to optimize discretionary expenses while healthy if retiree is still healthy at that year, or to use (3.3.2) to optimize withdrawals if the retiree is unhealthy at the current year. The GBM parameters of fund values are assumed constant through out the years.
Because of recalibration, for each testing scenario, we will have a sequence of withdrawals. To compare with the outputs from other models, we will take the average of the first 17 (HLE) years and then take the average of this value across all simulations. Following is the table showing initial withdrawals, average withdrawals, average ending fund value and the ruin probability.

<table>
<thead>
<tr>
<th>Model 4</th>
<th>Good</th>
<th>Moderate</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial withdrawals</td>
<td>$83,384</td>
<td>$83,384</td>
<td>$83,384</td>
</tr>
<tr>
<td>Average withdrawals for first 17 years</td>
<td>$88,294</td>
<td>$85,807</td>
<td>$83,050</td>
</tr>
<tr>
<td>Average ending fund value</td>
<td>$367,102</td>
<td>$331,387</td>
<td>$299,356</td>
</tr>
<tr>
<td>Ruin probability</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

We will also look at how the model actually performs in each simulated scenario. We first generate some paths of asset returns, as shown in Figure 4.3.3. We will then test how the model performs on each of these paths by assuming the retiree being healthy for 17 years (HLE) and then unhealthy for 5 years (ULE).

When market is good, the average withdrawal is $86,960 and the ending fund value when retiree dies is $412,232. Figure 4.3.4 shows how the fund value and withdrawals change over time.

When market is moderate, the average withdrawal is $79,450 and the ending fund value when retiree dies is $277,630. Figure 4.3.5 shows how the fund value and withdrawals change over time.

When market is bad, the average withdrawal is $78,182 and the ending fund value when retiree dies is $264,602. Figure 4.3.6 shows how the fund value and withdrawals change over time.
Figure 4.3.3: Return paths of three assets when market is good, moderate and bad
Figure 4.3.4: Fund values (left) and withdrawals (right) of Model 4 in good market

Figure 4.3.5: Fund values (left) and withdrawals (right) of Model 4 in moderate market

Figure 4.3.6: Fund values (left) and withdrawals (right) of Model 4 in bad market
4.3.5 Model 5: Kalman Filter Recalibration

This is our final model, which is Model 3 plus Kalman Filter recalibration. We created three Kalman Filter objects for three assets to dynamically adjust the estimation of the GBM parameters $\mu$’s and $\sigma$’s every year.

Same as Model 4, at the beginning of each year, the starting fund value and the age of retiree are changed as inputs to the optimal quantile solution. Also, we need to determine whether to use (3.3.1) to optimize discretionary expenses while healthy if retiree is still healthy at that year, or to use (3.3.2) to optimize withdrawals if the retiree is unhealthy at the current year. We will update the values of $\mu$’s and $\sigma$’s based on real asset daily returns from last year and use these values to simulate future returns and solve for the optimal quantile solution. This part has been explained in details in Figure 4.2.1.

Following is the table showing initial withdrawals, average withdrawals, average ending fund value and the ruin probability.

<table>
<thead>
<tr>
<th>Model 5</th>
<th>Good</th>
<th>Moderate</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial withdrawals</td>
<td>$83,384</td>
<td>$83,384</td>
<td>$83,384</td>
</tr>
<tr>
<td>Average withdrawals for first 17 years</td>
<td>$89,092</td>
<td>$86,027</td>
<td>$83,071</td>
</tr>
<tr>
<td>Average ending fund value</td>
<td>$340,089</td>
<td>$327,039</td>
<td>$314,041</td>
</tr>
<tr>
<td>Ruin probability</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

We will also look at how the model actually performs in each simulated scenario. We use the same generated paths shown in Figure 4.3.3 as in Model 4 and assume the retiree being healthy for 17 years (HLE) and then unhealthy for 5 years (ULE).

When market is good, the average withdrawal is $87,284 and the ending fund value when retiree dies is $363,198. Figure 4.3.7 shows how the fund value and withdrawals
4.3.6 Model Comparison

In general, the metrics we want to use to compare different models are ruin probability and average withdrawals. We want to achieve larger withdrawals while keeping ruin probability as small as possible. However, note that because of the different un-
derlying model assumptions, we sometimes can not say which model is better based on these metrics. But we will still try to draw insights by thinking this way.

Let’s first look at the three static models, Model 1-3.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
 & Model1 & Model2 & Model3 \\
\hline
Withdrawals & $82,815$ & $84,066$ & $86,180$ \\
\hline
\end{tabular}
\end{table}

We noticed that Model 2 has larger withdrawal than Model 1. Model 1 used the fixed death time at age 95 but Model 2 used random death time. In model 2, the LE is actually between 22 and 23, which means the expected death time is at around age 88. Thus there are many simulated scenarios in Model 2 when the retiree dies before age 95, which results in more possible annual withdrawals. Actually if we set death time to be age 85 in Model 1, then the optimal withdrawal of Model 1 would be larger than that of Model 2.

Withdrawal while healthy in Model 3 is larger than withdrawal in Model 2. This is also easy to understand because of the different underlying assumptions. In Model 3 we try to optimize the life style of a retiree while healthy and to withdraw less while unhealthy. Therefore, withdrawal while healthy in Model 3 must be larger than the constant withdrawal which we keep for the entire retirement life in Model 2.

Model 4 and 5 use annual recalibration and the benefits are prominent. Firstly,
the ruin probabilities calculated are 0%‘s in all six different testing set ups. Actually ruin could happen in more extreme scenarios (We will talk about it soon in this section), but it didn’t happen in the above case studies when we set the underlying $\mu$ of testing scenarios to be the original $\mu \pm 0.005$. The other advantage of recalibration is shown in the average ending fund values.

<table>
<thead>
<tr>
<th>Average ending fund value</th>
<th>Good</th>
<th>Moderate</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>$1,119,896</td>
<td>$671,162</td>
<td>$304,323</td>
</tr>
<tr>
<td>Model 2</td>
<td>$1,018,218</td>
<td>$730,340</td>
<td>$491,336</td>
</tr>
<tr>
<td>Model 3</td>
<td>$1,022,397</td>
<td>$772,747</td>
<td>$564,859</td>
</tr>
<tr>
<td>Model 4</td>
<td>$367,102</td>
<td>$331,387</td>
<td>$299,356</td>
</tr>
<tr>
<td>Model 5</td>
<td>$340,089</td>
<td>$327,039</td>
<td>$314,041</td>
</tr>
</tbody>
</table>

As we can see, Models with recalibration are much more flexible and the average ending fund value do not fluctuate much when market is good, moderate or bad. It means that with recalibration the models can allow retirees to spend more money when possible so as to enjoy retirement life to the full extent no matter the market conditions.

The results get more interesting when we compare Model 4 - Naive Recalibration with Model 5 - Kalman Filter Recalibration. Model 4 simply reruns Model 3 every year with different fund value and age of the retiree, while keeping the $\mu$’s and $\sigma$’s fixed when generating future simulations to solve for optimal quantile solution. Model 5 does it differently by estimating $\mu$’s and $\sigma$’s every year based on previous years’ daily asset returns and uses the newly estimated parameters to generate future simulations and solve for optimal quantile solution.
<table>
<thead>
<tr>
<th>Average withdrawals for first 17 years</th>
<th>Good</th>
<th>Moderate</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 4</td>
<td>$89,112</td>
<td>$81,387</td>
<td>$73,893</td>
</tr>
<tr>
<td>Model 5</td>
<td>$90,405</td>
<td>$82,305</td>
<td>$74,191</td>
</tr>
</tbody>
</table>

When market is good, Model 5 is able to know and is more positive about the future, thus outputs higher withdrawals. That is why Model 4 has fewer average withdrawals than Model 5 and has larger ending fund values.

When market is bad, Model 5 is also able to know and be more negative about the future, thus outputs lower withdrawals more actively and early than Model 4, which keeps withdrawing more money than it should and has to gradually decrease the withdrawals passively due to small fund value. Model 5 actually output smaller withdrawals at the first several years but our results show that in average Model 5 can output a little larger withdrawals for the first 17 years.

To our surprise, when market is moderate, even if Model 4 knows the original underlying \( \mu \)'s and \( \sigma \)'s, it still outputs fewer average withdrawals for the first 17 years than Model 5. This may because that for each scenario, Model 5 can adjust itself accordingly overtime and therefore achieve better overall performance.

Let us now look at more extreme cases when the asset returns get worse. We modify \( \mu \) this time by multiplying 0.2 and 0.1. We ran some fast tests based on 20 health simulations \( \times \) 20 fund value simulations and got the following results.

<table>
<thead>
<tr>
<th>Ruin</th>
<th>( \mu \times 0.2 )</th>
<th>( \mu \times 0.1 )</th>
<th>( \mu \times 0.05 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 4</td>
<td>2.5%</td>
<td>5%</td>
<td>6%</td>
</tr>
<tr>
<td>Model 5</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

We can see that ruin happened in our simulated scenarios when using Model 4 but still did not happen when using Model 5.
If we take the average withdrawals across all simulated scenarios for each year, we can get the path of averaged withdrawals for each year. Figure 4.3.10 shows the averaged withdrawals for Model 4 and Model 5 when we multiply $\mu$ by 0.2. We can see that Model 4 is not able to recognize a bad market and kept withdrawing too much in early years which resulting in a situation where it can only withdraw basic expenses and actually get ruined. Model 5 recognized a bad market only by one Kalman Filter update (after one year) and withdrew reasonable less money than Model 4, therefore is able to avoid ruin.

Actually we can easily make ruin happen for any single testing scenario when using Model 4 by increasing time horizon, increasing the ruin probability in quantile solution (e.g. from 5% to 10%) or just generating worse asset return paths. We omit these testing results here for brevity. In any case, the result is very promising for Model 5: ruin never happened unless we make the scenario unreasonably extreme.

To further demonstrate the fact that Kalman Filter recalibration can virtually eliminates ruin while market is bad and can help retirees withdraw as much as they can while market is good, let us look at the following tests. We keep the set-ups same as previous case studies except now we fix the health conditions of the retiree: she will...
get unhealthy at year 17 (age 82) and die at year 22 (age 87). Then we simulate 1,000 different scenarios of asset returns. We first calculate the optimal quantile solution based on the 1,000 simulated scenarios, which gives us $D = 77,794$ (Model 3). Thus the withdrawal is $D + B_h = 99,061$.

Since this solution is the 5% quantile, it means that there are 5% of the 1,000 (i.e., 50) scenarios where ruin happens if the retiree uses this solution. Now we apply Model 5 with Kalman Filter recalibration on these 50 scenarios and test the performances. We then look at the top 5% of the 1,000 scenarios where market condition is the best and using the initial withdrawal strategy $99,061 will result in largest remaining fund value at death time (year 22). We also apply Kalman Filter along each scenarios. The following are the comparison of the outputs. Withdrawals for Model 5 here are the average withdrawals while retiree is healthy (first 17 years).

<table>
<thead>
<tr>
<th>Worst 5% market scenarios</th>
<th>Ruin Probability</th>
<th>Withdrawals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 3 (static)</td>
<td>100%</td>
<td>$99,061</td>
</tr>
<tr>
<td>Model 5 (KF recalibration)</td>
<td>0%</td>
<td>$92,999, down by $6,062</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Best 5% market scenarios</th>
<th>Ruin Probability</th>
<th>Withdrawals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 3 (static)</td>
<td>0%</td>
<td>$99,061</td>
</tr>
<tr>
<td>Model 5 (KF recalibration)</td>
<td>0%</td>
<td>$106,053, up by $6,992</td>
</tr>
</tbody>
</table>

This simple test shows that Kalman Filter recalibration can eliminate ruin for the worst market conditions. Also, the average decrease in average withdrawals while healthy in the worst 5% market conditions is less than the average increase in withdrawals while healthy in the best 5% market conditions.

Based on all above analysis, we can conclude that model 5 beats all other models in the sense that it virtually eliminates ruin and can output largest possible suggested
withdrawals for retirees.

4.4 More Applications

4.4.1 Bequest Needs

Retirees may have the needs to determine how much money can be left after enjoying their own lives. If the retiree has certain bequest needs, we can add it into our model by simply modifying the closed-form solution (3.3.1) and (3.3.2). There, instead of letting $T_{LR} = 0$, we now set $T_{LR} = \text{”bequest need”}$ to solve for $D$. But there may be a problem if the retiree sets the bequest need too high. In this case, the retiree will get ruined during retirement and could not withdraw any money due to a too high bequest requirement. Therefore, similar to basic expenses, we can give a suggested bequest amount. We can use the average value over all ending values of retirement Fund. For example, a suggested bequest needs in our case study when market is bad would be $327,039.

4.4.2 Liability Cash flow for Fund Manager

In this section, we will look at how our model can help a fund manager doing Asset and Liability Management (ALM) by providing the predicted future liability cash flow.

Let us denote $L_i$ the liability at year $i$, which is how much the fund manager needs to pay the retiree each year. In our model, this depends on whether the retiree is healthy, unhealthy or dead. Denote the DIA payout as $DIA_i = A_d \cdot p_d \cdot \mathbb{I}_{i \geq y_{DLA}}$. If
the retiree is healthy, the cash flow from fund to retiree is \( CF_i^H = B_h + D_h - C - DIA_i \), which is basic expenses while healthy plus discretionary expenses while healthy, minus social security income, then minus the DIA payout if year \( i \) is after the predetermined DIA payout date. If the retiree is unhealthy, the fund needs to give retiree \( CF_i^U = B_u - C - DIA_i \), which is basic expenses while unhealthy minus social security income, then minus the DIA payout. If the retiree dies at year \( i \), the fund manager needs to pay all the remaining fund value to the retiree, which is \( CF_i^D = T_i \) in our notation.

Therefore in summary, the expected liability at year \( i \) is

\[
L_i = p_i^H CF_i^H + p_i^U CF_i^U + p_i^D CF_i^D,
\]

where \( CF_i^H, CF_i^U, CF_i^D \) are defined above and \( p_i^H, p_i^U, p_i^D \) are the probabilities of the retiree being healthy, unhealthy and dead at year \( i \).

We can run simulations to determine the probabilities.

\[
L_i = \frac{1}{|S|} \sum_{s \in S} I_i^H(s)CF_i^H + I_i^U(s)CF_i^U + I_i^D(s)CF_i^D,
\]

where \( I \)'s are the indicator of healthy conditions for simulation \( s \) in the set of all simulations \( S \). For example, \( I_i^H(s) = 1 \) if retiree is healthy in year \( i \) for simulation \( s \) and \( I_i^H(s) = 0 \) if retiree is unhealthy or dead in year \( i \) for simulation \( s \).

We will follow the case study in Section 4.3. When market is moderate, the curve in Figure 4.4.1 shows the expected future liability cash flow at the beginning year (when retiree is at age 65). Because of annual recalibration with Kalman Filter, our model can update the expected future liability cash flow every year. Figure 4.4.2 shows the updated liability cash flow for the first 3 years and first 5 years. Note that there is a steep drop at year 17 due to the predicted change in health status.
Figure 4.4.1: Liability cashflow predicted at year 0 (age 65)

Figure 4.4.2: Recalibrated liability cashflows for the first 3 years and 5 years.
Bibliography


