PRICING A SEQUENCE RISK PROTECTION PRODUCT

Authors:

Jack Cascone, Matthew Letourneau, and Zachary Pitts

Advisors:

Barry Posterro, Jon Abraham

April 25, 2024

Worcester Polytechnic Institute

Major Qualifying Project

Abstract

If an investment account grows based on variable returns, such as those from the S&P500, they may be vulnerable to sequence risk. Sequence risk is the volatility of the ultimate account balance of an investment involving multiple deposits that arises due to the possible permutations of annual returns. Two sequences of returns with the same annualized average can result in different final account balances if the investor makes multiple deposits across time. Our group sought to create a product that protects against the effects of sequence risk in an investment account that follows the S&P500 over a 30-year period, where the customer makes identical annual deposits. We used simulation techniques to investigate the premium needed for this protection at different confidence levels. This paper discusses those efforts and the various challenges that come with pricing and diversifying this policy.

Table of Contents

Background

Sequence Risk

Financial awareness has reached a new height in popularity. Many people, regardless of career or background, are looking to make investments that will set them up for success in the future. The stock market's capability to generate massive returns quickly is enticing. However, the risk of losing everything is too large for some investors. Nevertheless, profitable investments are in the front of everyone's mind. To establish the background of our report, we must first discuss the concept that is the basis of our goal to combine finance and insurance: sequence risk. Interest rates are constantly changing, especially within the stock market. It is inevitable that there will be periods of time when an investment either underperforms or loses money. Sequence risk is "the risk of experiencing bad investment outcomes at the wrong time" (Clare et al., 2017). In other words, there are opportune and inopportune times for poor performance of an investment. Depending on the ordering of interest rates, an investment will have varying outcomes. This can be demonstrated through a few simple examples.

Basic Examples

Before we begin with our examples, there are a few things that are important to note. For sequence risk to exist, an investment must be made at incremental time periods. A single investment does not have any sequence risk. Additionally, sequence risk and the length of the investment plan have a direct relationship. This is because for every permutation of interest rates, a larger time frame includes that permutation and more. To demonstrate this point, consider the following scenario. An investment of 1,000 dollars is made at the beginning of year one. This investment grows each year at a variable interest rate. There are two possibilities: The rates increase by one percent every year from 1% to 5% or decrease every year from 5% to 1%.

Which option is better? The answer is neither! Both accounts will have the same final account value. This can be seen in the chart below where account "A1" grows by the increasing rates, and account "A2" grows by the decreasing rates.

Figure 1: Example table of a lump sum investment without sequence risk

	Account Value ($)				
Year	1	2	3	4	5
A1	1010.00	1030.20	1061.11	1103.55	1158.73
A2	1050.00	1092.00	1124.76	1147.26	1158.73
Difference	-40.00	-61.80	-63.65	-43.70	0.00

As seen in the table, the difference between the final balance in each account is $0. As previously mentioned, this is because a single investment does not hold sequence risk. Now, we can begin our further examples.

For this first example we will consider the scenario where an investment of 1,000 dollars is made at the beginning of every year for five years. For simplicity, there will not be a different interest rate every year. Instead, all but one year will yield a nine percent return and there will be an eight percent return for the remaining year. The figure below depicts each permutation (P#) of growth factors for this example, where each column represents a new year, and each row represents the growth of the account over the five years.

Figure 2: Table of growth factors with one lower value, creating sequence risk

	Growth Factors				
Year	1	2	3	4	5
P1	1.08	1.09	1.09	1.09	1.09
P2	1.09	1.08	1.09	1.09	1.09
P3	1.09	1.09	1.08	1.09	1.09
P4	1.09	1.09	1.09	1.08	1.09
P5	1.09	1.09	1.09	1.09	1.08

When there is a single lower growth factor, it is best for it to occur when the least amount of money is invested and worst to have it when the most amount of money is invested. In permutation one (P1) we get the most money in the account after five years, despite having the least amount of money after one year. Conversely, P5 has the largest account the first four years but becomes the smallest balance after the fifth. Each of the other permutations follow this same pattern. The earlier the year with an eight percent interest rate, the higher the final account value. This can be seen in the figure below.

Figure 3: Example table of an annual investment with sequence risk

	Account Value ($)					Difference
Year	1	2	3	4	5	From Geo Avg
P1	1080.00	2267.20	3561.25	4971.76	6509.22	23.89
P2	1090.00	2257.20	3550.35	4959.88	6496.27	10.94
Geo Avg	1087.99	2271.72	3559.61	4960.82	6485.33	0.00
P3	1090.00	2278.10	3540.35	4948.98	6484.39	-0.94
P4	1090.00	2278.10	3573.13	4938.98	6473.49	-11.84
P5	1090.00	2278.10	3573.13	4984.71	6463.49	-21.84

In the figure above, the central row, labeled "Geo Avg," with the final column in orange was generated to grow following the geometric average of the interest rates for the five years: 1.088. Each account value is colored in either green or red, if their balance is above or below the geometric average at that time, respectively. It is important to note that each of these accounts are making money in this example. Being in the red does not necessarily mean loss and being in the green does not necessarily mean profit. We are simply comparing the account values to the geometric average rate. This comparison at the end of the five years can be seen in the final column, "Difference From Geo Avg."

For the next example, we will demonstrate that for any set of interest rates, the accounts will still follow this pattern and that the accounts with a favorable sequence, will finish above the account that follows the geometric average growth factor, regardless of the actual return values. For this example, we will again invest 1,000 dollars at the beginning of each year. This time, however, we will only invest for four years, and our interest rates will be -2%, -1%, 1%, and 2%. There are 24 possible permutations of these rates, and this can be seen as P1 through P24 in the tables below.

Figure 4: Table of all possible permutations of 4 growth factors

	Growth Factors			
Year	1	2	3	4
P1	0.98	0.99	1.01	1.02
P2	0.98	0.99	1.02	1.01
P3	0.98	1.01	0.99	1.02
P4	0.98	1.01	1.02	0.99
P5	0.98	1.02	0.99	1.01
P6	0.98	1.02	1.01	0.99
P7	0.99	0.98	1.01	1.02
P8	0.99	0.98	1.02	1.01
P9	0.99	1.01	0.98	1.02
P10	0.99	1.01	1.02	0.98
P11	0.99	1.02	0.98	1.01
P12	0.99	1.02	1.01	0.98

	Growth Factors			
Year	1	2	3	4
P13	1.01	0.98	0.99	1.02
P14	1.01	0.98	1.02	0.99
P15	1.01	0.99	0.98	1.02
P16	1.01	0.99	1.02	0.98
P17	1.01	1.02	0.98	0.99
P18	1.01	1.02	0.99	0.98
P19	1.02	0.98	0.99	1.01
P20	1.02	0.98	1.01	0.99
P21	1.02	0.99	0.98	1.01
P22	1.02	0.99	1.01	0.98
P23	1.02	1.01	0.98	0.99
P24	1.02	1.01	0.99	0.98

Based on these permutations, we have again created a table which shows the account balance at the end of each year. Figure 5 follows the same format as the previous example and can be seen below.

Figure 5: Table of all possible permutations as policies, ordered by difference from Geo Avg

	Account Value ($)				Difference
Year	1	2	3	4	From Geo Avg
P1	980.00	1960.20	2989.80	4069.60	70.85
P2	980.00	1960.20	3019.40	4059.60	60.85
P7	990.00	1950.20	2979.70	4059.30	60.55
P8	990.00	1950.20	3009.20	4049.30	50.55
P3	980.00	1999.80	2969.80	4049.20	50.45
P5	980.00	2019.60	2989.40	4029.30	30.55
P9	990.00	2009.90	2949.70	4028.70	29.95
P4	980.00	1999.80	3059.80	4019.20	20.45
P13	1010.00	1969.80	2940.10	4018.90	20.15
P6	980.00	2019.60	3049.80	4009.30	10.55
P11	990.00	2029.80	2969.20	4008.90	10.15
P15	1010.00	1989.90	2930.10	4008.70	9.95
Geo Avg	999.87	1999.62	2999.25	3998.75	0.00
P19	1020.00	1979.60	2949.80	3989.30	-9.45
P14	1010.00	1969.80	3029.20	3988.90	-9.85
P10	990.00	2009.90	3070.10	3988.70	-10.05
P21	1020.00	1999.80	2939.80	3979.20	-19.55
P12	990.00	2029.80	3060.10	3978.90	-19.85
P20	1020.00	1979.60	3009.40	3969.30	-29.45
P16	1010.00	1989.90	3049.70	3968.70	-30.05
P17	1010.00	2050.20	2989.20	3949.30	-49.45
P22	1020.00	1999.80	3029.80	3949.20	-49.55
P23	1020.00	2040.20	2979.40	3939.60	-59.15
P18	1010.00	2050.20	3019.70	3939.30	-59.45
P24	1020.00	2040.20	3009.80	3929.60	-69.15

Here, our geometric average growth factor is 0.999, much lower than the previous example. This results in a final Geo Avg account balance of $3,998.75. However, as shown in the final column, half of the permutations still result in accounts that are in the green. This is the strength of sequence risk that we hope to harness as a potential product. The interest rates themselves do not affect the relation to the geometric average. The real importance is in their order. In the permutations with a favorable order, the accounts are still above the geometric average, even though they did not have as much growth as in the previous examples. There is still a large risk in this example because the geometric standard deviation is large. The largest possible deficit from the geometric average is \$69.15, which is much greater than in the previous example, but the highest possible surplus is also much greater, at \$70.85.

In this next example, we want to demonstrate the scenario where the geometric average is the same as the previous example, but the geometric standard deviation is changed. In this example, we have doubled the geometric standard deviation and kept the mean constant. The new rates can be seen in Figure 6 below and follow the same form as the previous example.

Figure 6: Table of all possible permutations of 4 growth factors, with larger standard deviation

	Growth Factors					Growth Factors			
Year	1	2	3	4	Year	1	2	3	4
P1	0.96052	0.980223	1.02023	1.04053	P13	1.02023	0.96052	0.980223	1.04053
P2	0.96052	0.980223	1.04053	1.02023	P14	1.02023	0.96052	1.04053	0.980223
P3	0.96052	1.02023	0.980223	1.04053	P15	1.02023	0.980223	0.96052	1.04053
P4	0.96052	1.02023	1.04053	0.980223	P16	1.02023	0.980223	1.04053	0.96052
P5	0.96052	1.04053	0.980223	1.02023	P17	1.02023	1.04053	0.96052	0.980223
P6	0.96052	1.04053	1.02023	0.980223	P18	1.02023	1.04053	0.980223	0.96052
P7	0.980223	0.96052	1.02023	1.04053	P19	1.04053	0.96052	0.980223	1.02023
P8	0.980223	0.96052	1.04053	1.02023	P20	1.04053	0.96052	1.02023	0.980223
P9	0.980223	1.02023	0.96052	1.04053	P21	1.04053	0.980223	0.96052	1.02023
P10	0.980223	1.02023	1.04053	0.96052	P22	1.04053	0.980223	1.02023	0.96052
P11	0.980223	1.04053	0.96052	1.02023	P23	1.04053	1.02023	0.96052	0.980223
P12	0.980223	1.04053	1.02023	0.96052	P24	1.04053	1.02023	0.980223	0.96052

Again, we have the same table as in the previous two examples, depicting the account balance at the end of each year, as well as the difference from the geometric average.

Figure 7: Table of all possible permutations, ordered by difference from Geo Avg, with larger standard deviation

	Account Value ($)				Difference
Year	1	2	3	4	From Geo Avg
P1	960.52	1921.75	2980.85	4142.19	143.44
P2	960.52	1921.75	3040.16	4121.89	123.14
P7	980.22	1902.04	2960.74	4121.27	122.52
P8	980.22	1902.04	3019.66	4100.97	102.22
P3	960.52	2000.18	2940.84	4100.56	101.81
P5	960.52	2039.98	2979.86	4060.36	61.61
P9	980.22	2020.28	2901.04	4059.15	60.40
P4	960.52	2000.18	3121.77	4040.26	41.51
P13	1020.23	1940.47	2882.31	4039.66	40.91
P6	960.52	2039.98	3101.47	4020.35	21.60
P11	980.22	2060.48	2939.65	4019.34	20.59
P15	1020.23	1980.27	2862.61	4019.16	20.41
Geo Avg	999.87	1999.62	2999.25	3998.75	0.00
P19	1040.53	1959.97	2901.43	3980.35	-18.40
P14	1020.23	1940.47	3059.65	3979.36	-19.39
P10	980.22	2020.28	3142.69	3979.14	-19.61
P21	1040.53	2000.17	2881.73	3960.24	-38.51
P12	980.22	2060.48	3122.39	3959.64	-39.11
P20	1040.53	1959.97	3019.84	3940.34	-58.41
P16	1020.23	1980.27	3101.06	3939.15	-59.60
P17	1020.23	2102.11	2979.64	3900.93	-97.82
P22	1040.53	2000.17	3060.86	3900.54	-98.21
P23	1040.53	2081.80	2960.14	3881.81	-116.94
P18	1020.23	2102.11	3040.76	3881.23	-117.52
P24	1040.53	2081.80	3020.85	3862.11	-136.64

Here, we still have the same final account balance of from the Geo Avg row, $3998.75. We also have the same 12 accounts in the green and the same 12 accounts in the red. The difference in this example is the magnitude of the values in the final column. The best and worst possible differences have both nearly doubled. Although the deficits are worse in this example, the surpluses are inversely better. This is the effect of changing the geometric standard deviation of the interest rates while maintaining the geometric average.

Product Introduction

Let's look at what might happen when making investments in the stock market, which inherently has random returns. Specifically, we will be making $1,000 investments into the S&P500 every year, for 30 years. Because of the random returns in the stock market, this investment strategy has sequence risk. What if we wanted to invest in the S&P without sequence risk? As shown in Figure 1, if we made one lump sum investment at the beginning, the final account balance will be the same regardless of the order of the growth factors. What if we are not financially capable of making this large lump sum investment, but we still want to avoid any downside from investing with sequence risk. From the Basic Examples section, recall that if we have different growth rates, the order of them creates sequence risks. Therefore, for this investment pattern to be free of sequence risk, our growth factors need to be constant. If this constant rate is equal to the geometric average, then the account balance will be equal between the lump sum investment and the annual investment. If investments are constant and annual, this holds true no matter the size of our investments, though the lump sum deposit that is equivalent will be altered.

This project attempts to create an insurance product where you can make annual investments in the S&P500 and be protected from sequence risk exposure. The final account balance where the growth factors are equal to the geometric average every year (The "Geo Avg" rows in our basic examples) is known as the "strike price". If the final balance of one of policyholders is less than the strike price, we would pay them the difference. This insurance product is very similar to a European style put option.

The specifics for our product are 30 annual investments of $1,000 with the random growth rates of the S&P500, we guarantee their investment will be at least equal to the strike price calculated from the geometric average, and we pay out the difference if their final balance falls short of that guaranteed amount. It is worth noting that this strike price is not known until the end of the final investment year.

Risk Mitigation: Diversification Vs. Hedging

There is a fundamental difference between insurance products and finance products that needs to be clear for the product that we are creating. In insurance, the expected value of profit for any single policy is positive. Of course, losses are inevitable, and some policies are bound to lose money. However, the more policies that the insurance company sells, the more likely they are to achieve their expected profit. This is called the law of large numbers and because of this, insurance companies have an incentive to diversify their policies. For example, if an insurer has policies that cover losses to natural disasters, it would be extremely risky to have all their policies in a single coastal town, where one hurricane could cause losses to all of them. It would be much better to have many policies spread out so that in the case of a loss, most policies are unaffected by the same event, this is called diversification. Insurance products calculate the probability of any event happening and set a price accordingly. Insurance products profit from expected values and the law of large numbers, whereas finance products profit from trends in interest rates.

For a finance product with a variable interest rate to be profitable, interest rates or stock prices need to make a favorable change after the investment has been made. If someone who sells these products were to employ the same insurance strategy of diversification to their products, they would be taking a massive risk. For products that invest in the stock market or try to predict interest rates, selling more and more policies multiplies the risk. Since there can only be one realization of the stock market and interest rates, every investor has the same outcome. So, if there is a loss, more investors mean more losses. Because of this, a new strategy must be used to mitigate risk. The strategy used for these types of products is called hedging. Hedging is when there is an existing risk, purchasing an opposite risk to balance it. For example, in the stock market, large tech companies and large banks often have inverse outcomes. When one increases, the other decreases. Thus, it makes sense to invest in both large tech companies and large banks. This way, the money grows with the general rise of the market instead of gambling on which sector is going to have a boom.

For the product that we are creating, although it has its roots in insurance, it is essentially a financial product. We will attempt to find methods of diversification, but we predict that the best option for mitigating risk in this scenario is hedging. This is not something that is in the scope of our project but is something that can be done in future projects and must be done if this product is to go onto the market.

Geometric Statistics

When dealing with compounding values such as interest rates, it is important to use summary statistics that accurately reflect the behavior of the interest rates on the total. For example, if we have two annual growth rates of 50% and 150%, their arithmetic average (the common meaning of "average") is 100%. However, growing a sum of $1000 through rates of 50% and 150% will result in a final value of $750. Intuitively, the "average" growth for compounding numbers in this context should really be less than 100%, to reflect this property of compounding numbers. This average is referred to as the "geometric average", and is calculated with the following formula for a set of values $A = \{a_1, a_2, \ldots, a_n\}$:

$$Geo(A) = Product(a_1, a_2, \ldots, a_n)^{\frac{1}{n}}$$

With the example growth rates of 50% and 150%, this results in a geometric average of 86.6%, which accurately reflects the final total of $750 after 2 years of growth. This is the process that we use in calculating our strike price.

$$Geo(A) = Product(0.5,1.5)^{\frac{1}{2}}$$

$$= (0.75)^{\frac{1}{2}} = .866\ldots$$

When working with a geometric average, other formulas and calculations that would typically use an arithmetic average need to be altered. One important example of this includes standard deviation. Below is the formula for calculating the geometric standard deviation, GeoSD(A), for a set of numbers A = $\{a_1, a_2, \ldots, a_n\}$, with geometric mean Geo(A), is:

$$GeoSD(A) = \exp\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\ln\left(\frac{a_i}{Geo(A)}\right)^2\right)}\right)$$

In simpler terms, this formula takes the natural logarithm of each value divided by the geometric average, calculates the arithmetic standard deviation, then exponentiates the result.

Changing the standard deviation of a set without changing the average

The following section was important to our methods but is not necessary to understand our analysis or conclusions.

To test out how the geometric standard deviation of growth factors impacts sequence risk, we needed a method of altering the spread of a set of values, without changing their geometric average. Our method involves performing the following operations on each initial growth factor a_i:

$$A = \left\{a_1, a_2, \dots a_n\right\}$$

$$B = \left\{b_1, b_2, \dots b_n\right\}$$

$$b_i = \left(\frac{a_i}{Geo(A)}\right)^c \cdot Geo(A)$$

This process will result in a set, B, with the same geometric average, but a geometric standard deviation raised to the power c. The proof of the first of these two properties of this process is shown on the next page.

Given sets A and B defined above, we'll start by writing out the formula for each of their geometric averages:

$$Geo(A) = \left(\prod_{i=1}^{n} a_i \right)^{\frac{1}{n}}$$

$$Geo(B) = \left(\prod_{i=1}^{n} b_i \right)^{\frac{1}{n}}$$

We can substitute b_i for its stated definition in terms of a_i:

$$Geo(B) = \left(\prod_{i=1}^{n} b_i \right)^{\frac{1}{n}}$$

$$= \left(\prod_{i=1}^{n} \left(\frac{a_i}{Geo(A)} \right)^{c} \cdot Geo(A) \right)^{\frac{1}{n}}$$

Since Geo(A) and 'c' are both constants, we can pull them both out of the product and simplify in the following steps:

$$Geo(B) = \left(\prod_{i=1}^{n} \left(\frac{a_i}{Geo(A)} \right)^c \right)^{\frac{1}{n}} \left(Geo(A)^n \right)^{\frac{1}{n}}$$

$$= \left(\prod_{i=1}^{n} \left(\frac{a_i}{Geo(A)} \right)^c \right)^{\frac{1}{n}} \cdot Geo(A)$$

$$= \left(\left(\prod_{i=1}^{n} \left(\frac{a_i}{Geo(A)} \right) \right)^{\frac{1}{n}} \right)^c \cdot Geo(A)$$

$$= \left(\left(\prod_{i=1}^{n} (a_i) \right)^{\frac{1}{n}} \right)^c \cdot \left(\left(Geo(A)^{-n} \right)^{\frac{1}{n}} \right)^c \cdot Geo(A)$$

$$= \left(\left(\prod_{i=1}^{n} (a_i) \right)^{\frac{1}{n}} \right)^c \cdot Geo(A)^{-c} \cdot Geo(A)$$

:

Notice that within this formula, we can see the exact definition for Geo(A) stated earlier, so we can substitute that in and make some cancellations to reach our final value:

$$Geo(B) = Geo(A)^c \cdot Geo(A)^{-c} \cdot Geo(A)$$

$$= Geo(A)$$

To better visualize what is going on here, let's look at an example performing this method on a set of 5 growth rates.

Figure 8: Example table of growth factors before (Set A) and after (Set B) using the method

Set A	Values		Set B
a_1	110%	118.87%	b_1
a_2	120%	141.46%	b_2
a_3	90%	79.57%	b_3
a_4	80%	62.87%	b_4
a_5	115%	129.92%	b_5
Geo(A)	1.017937	1.017937	Geo(B)
GeoSD(A)	1.190114	1.416371	GeoSD(B)

Notice here that Geo(A) = Geo(B) but GeoSD(B) is now much larger than GeoSD(A).

The proof for how the geometric standard deviation is altered here is in Appendix A.

Findings and Analysis

Building the Simulation

To estimate returns from future investments in the S&P500, we conducted simulations using historical data. We constructed the list of annual returns, randomized their order, and selected the first 30 annual returns. The geometric average of these 30 annual returns is 1.076 and the standard deviation of these returns is 0.205, both rounded to the nearest thousandth. As seen below, 17 of these rates fall above the geometric average and 13 fall below it.

Figure 9: The 30 Annual Growth Factors Collected for Simulation

1.140	1.122	1.131	0.943	0.898	1.089
1.161	1.031	1.166	1.620	0.926	1.011
0.902	1.110	1.061	1.156	1.522	1.116
0.680	1.045	1.114	1.028	1.127	1.098
1.139	0.794	0.887	0.973	1.206	1.610

Establishing the Investment Plan

Using the data from the S&P500, we generated a simulation consisting of 10,000 unique permutations of the 30 annual returns. In this simulation, each sequence of returns corresponds to a single policyholder. Their policy will require them to invest $1000 into the S&P500 at the beginning of each year, for 30 years. This plan allows the simulation to demonstrate the full effect of sequence risk, as each new interest rate is met with an additional investment.

Each year, the money in the account grows at that years' interest rate, and after 30 years, the final account balance is reached. At this point, we can calculate the strike price. Recall that the strike price is the final account balance if the account's growth factor was the geometric average rate determined from the price yield of the S&P500. As stated, the geometric average of the 30 returns is about 1.076. At the investment pattern for this simulation, this rate results in a final strike price of $113,241. Based on this strike price, we can calculate the payment that we would owe each policyholder. If an individual account is lower than the strike price, we would pay the difference. If an account is above the strike price, the policyholder would keep their balance, and we would record a payment of $0. It is important to remember that in a real scenario, each policyholder would experience the same rates and would have the same payout, but this allows us to analyze each possibility.

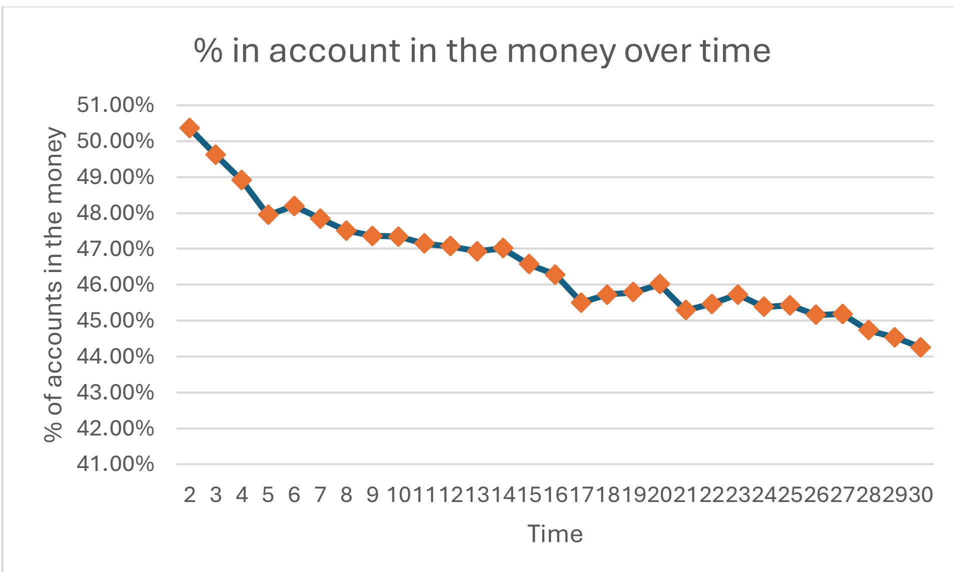
The average payment from the 10,000 permutations in this simulation is $8,476. We can also calculate the average of the payments for the policies excluding those with a payment of $0. This average is $19,154, which is sensibly much higher, since it is all possible scenarios given that there is a payment.

Risk Analysis

After conducting the simulation, there are some important points of analysis to note. We will begin with which accounts are "in the money." An account "in the money" is below the strike price, and thus we owe them a payment. Of the 10,000 permutations, 4,426 of them resulted in accounts in the money. Equivalently, we make payments 44.26% of the time.

While each account will have the same geometric average at the end of 30 years, each will have its own running geometric average to date at each point throughout the policy period. This only considers the return rates that the account has experienced so far. We can look at the percentage of accounts that are in the money at any given time relative to their own current geometric average. At time 1, everyone's account balance is exactly equal to their running geometric average balance since the rates are the same. Starting at time 2 we have 50.35% of accounts in the money. The percentage of accounts in the money almost always decreases year over year. At the end of 30 years, we reach our lowest percentage of accounts in the money at 44.26%.

Figure 10: Graph of percentage of accounts in the money over time



The growth rates for our simulation include three rates that are noticeably higher than the others. There are four key sequences that must be considered to demonstrate the impact that the placement of the high growth rates has on the final account balance. These sequences are the best and worst possible permutations, and the best and worst permutations that occurred in the simulation, which we refer to as best and worst realized. The best possible scenario is where the interest rates are strictly increasing, meaning the three high rates are at times 28, 29, and 30. This permutation results in a final account balance of $467,890. The worst possible scenario is where the interest rates are strictly decreasing, so the high rates are times 1, 2, and 3. This permutation results in a final account balance of $37,843. Compared to the strike price, $113,241. The best scenario is 4.13 times greater than this, and the worst scenario is 2.99 times smaller than this. This is initially a good indication that this product could be successful, the reward is greater than the risk relative to the strike price. Below are graphs of the returns from the best and worst realized scenarios as well as best and worst possible scenarios. The gray markers are indicative of the three high rates, and the orange bar is the geometric average.

Figure 11: Graph depicting the order of growth factors in the worst simulated case

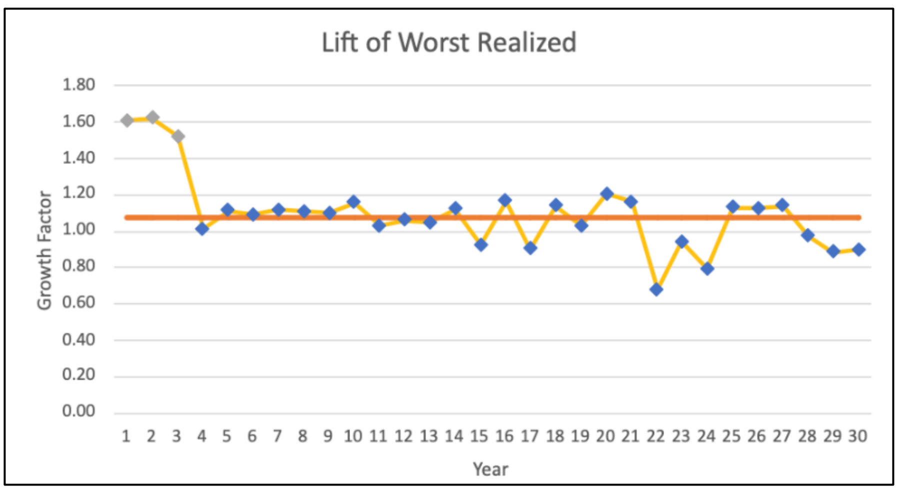


Figure 12: Graph depicting the order of growth factors in the worst possible case

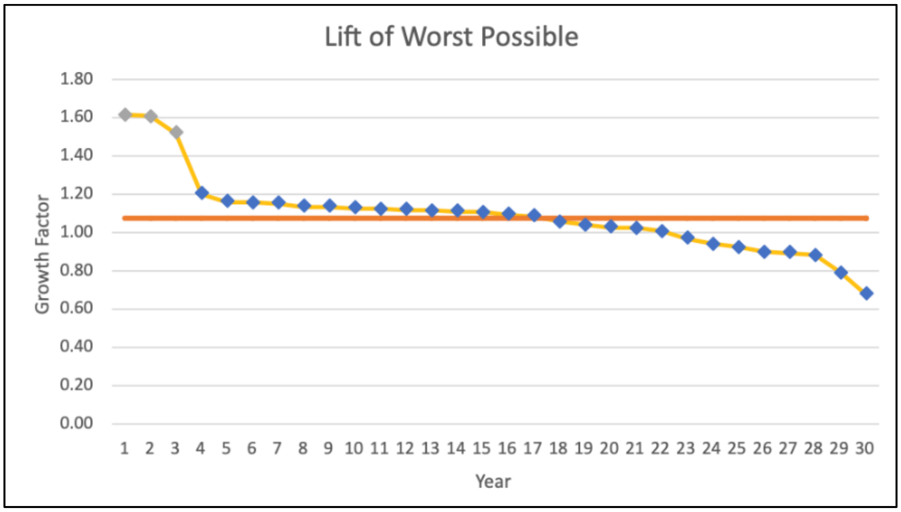


Figure 13: Graph depicting the order of growth factors in the best simulated case

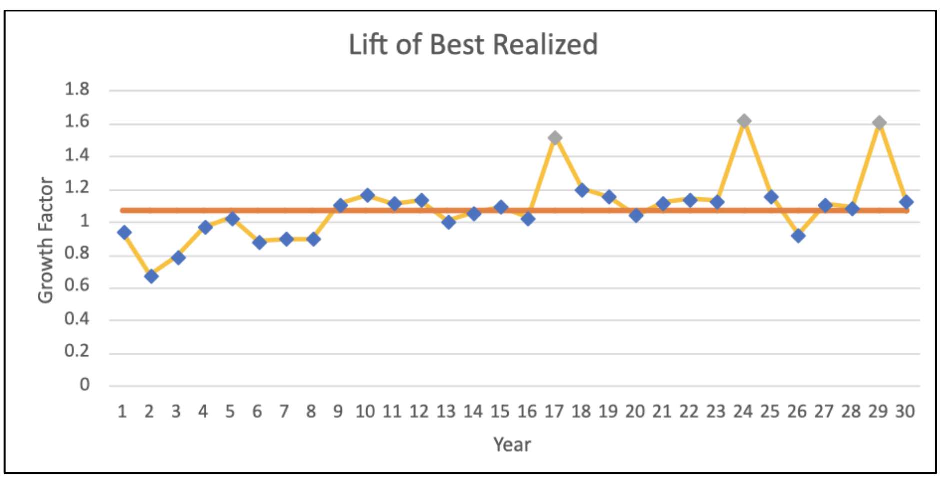
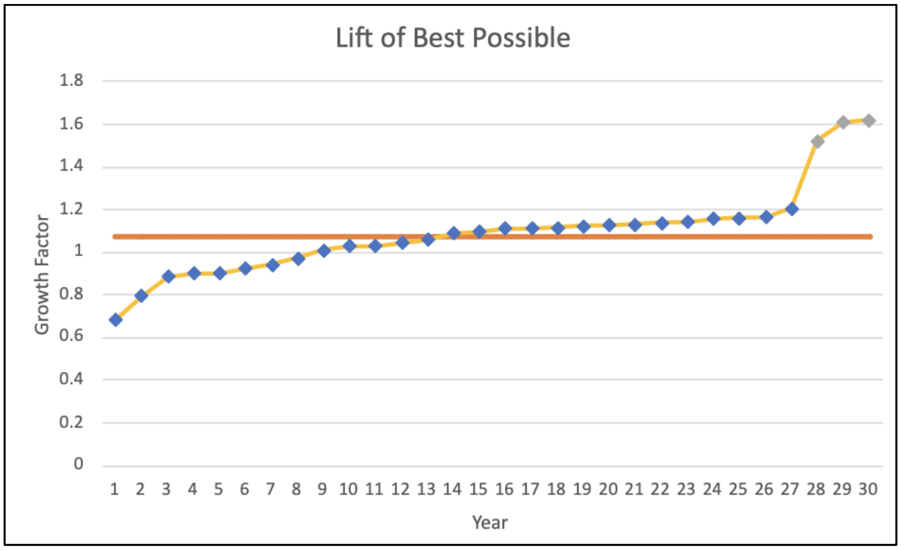


Figure 14: Graph depicting the order of growth factors in the best possible case



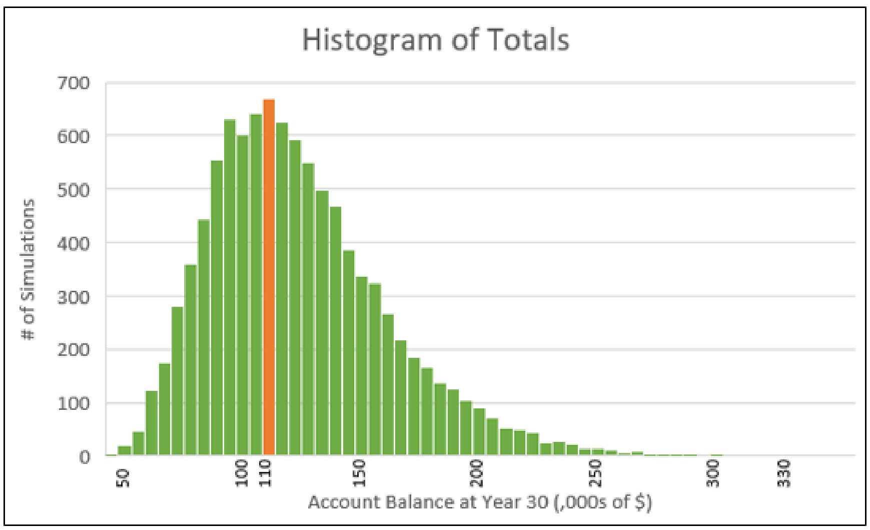
The best realized permutation has the high rates occurring at times 17, 24, and 29. While this is noticeably off from their ideal placements at the end, they still all occur in the second half of the investment period, and most of the other rates that are above the geometric average still occur in the later period. The final account balance of this permutation is $327,722, which is 2.91 times greater than the strike price, and $138,169 less than the best possible balance. This difference can be mostly attributed to the less-than-ideal positions of the three high rates.

The worst realized permutation has these high rates also occurring at times 1, 2, and 3, though not quite in descending order. While some of the other rates above the average do occur at the beginning as well, the remaining rates are from in descending order. The final balance from this permutation is $49,130, which is 2.304 times smaller than the strike price and $11,286 more than the worst possible permutation. Because the high rates are a lot closer to their placement in the worst possible permutation, the final account balance is much more comparable to it than the best realized is comparable to the best possible.

Distribution of Simulations

A core part of our analysis from conducting the simulations was to understand the distributions of account balances and payments. This was meant to give us an estimate of the magnitude of sequence risk for this set of 30 growth factors.

Figure 15: Histogram of the distribution of final account balances from simulations



This histogram depicts the account balances of the 10,000 simulations. The orange bin represents the horizontal location of the strike price, which is this graph's mode. However, the distribution is skewed right, appearing to represent a lognormal distribution.

Figure 16: Histogram of the distribution of final account balances from simulations using a lower standard deviation of growth factors

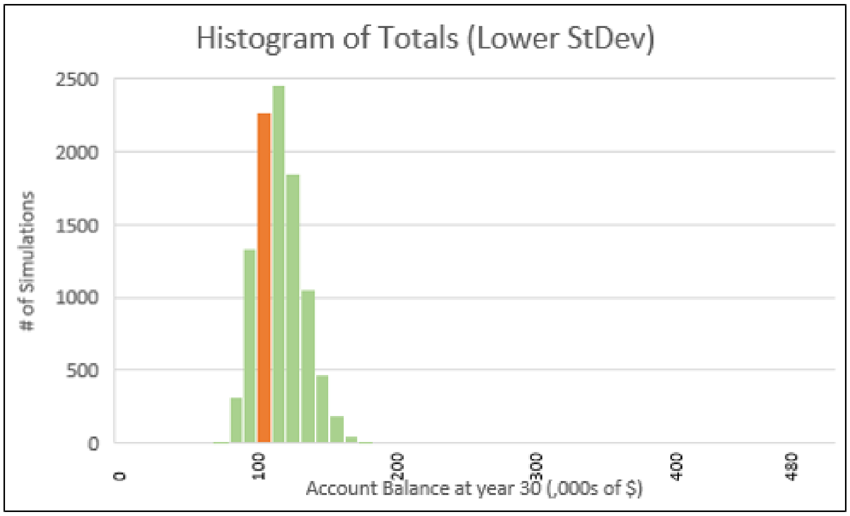
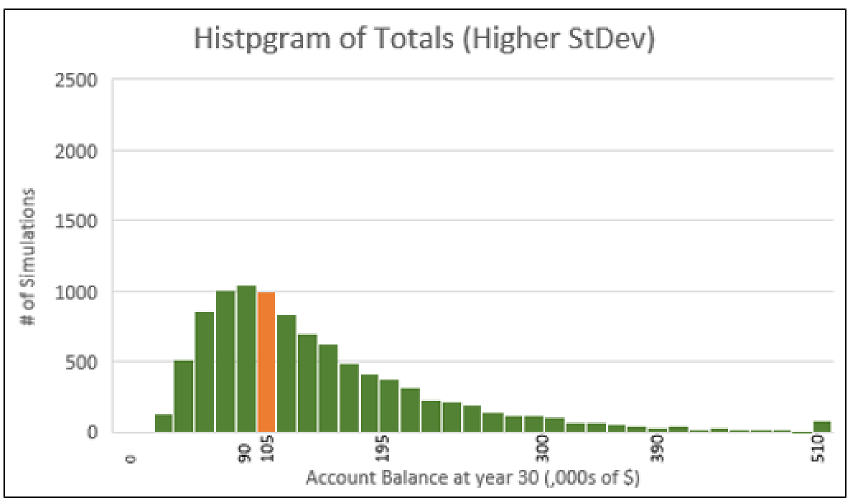


Figure 17: Histogram of the distribution of final account balances from simulations using a higher standard deviation of growth factors



The two histograms above depict two additional sets of 10,000 simulations each with higher and lower standard deviations of the growth factors (approximately 0.5 and 2.0 times the standard deviation for each set of simulations respectively). These sets of growth factors were scaled while keeping the geometric average using the scaling method detailed in the background. Each histogram uses the same axis boundaries for visual comparison.

Intuitively, the simulation with a higher standard deviation across the input growth factors resulted in a more spread-out distribution for the output account balances. The lower standard deviation input resulted in an average account balance of $115,418, and an average payment of $5,216. The higher standard deviation input resulted in an average account balance of $153,488, and an average payment of $12,938. The higher deviation input resulted in both a higher average account balance and a higher payment, even though the strike price remained the same. This is because the far-right tail values on the higher deviation input end up bringing the average account balance much higher, but those do not impact the average payment since those accounts are not in the money.

Pricing Strategy

Now that we have finished conducting our simulations and analyzing the outcomes, we must begin to calculate the premiums that we would charge for this product. In a perfect world, the strike price will always be below the final account balances. In reality, we know that this will not be the case if this product goes to the market. There will eventually be losses, and we need to adequately charge premiums that will cover them.

Black-Scholes Method

Economists Fisher Black and Myron Scholes explored creating risk free portfolios. In an article titled "The Pricing of Options and Corporate Liabilities" they shared the formula now known as the Black-Sholes Formula which can be used to price different stock market options. Our product is like a put option, as our product allows the customer to put their investment with us, which the Black-Scholes Formula can price.

Formula

$$C = N(d_1)S_t - N(d_2)Ke^{-rt}$$

$$\text{where } d_1 = \frac{\ln \frac{S_t}{K} + (r + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}$$

$$\text{and } d_2 = d_1 - \sigma\sqrt{t}$$

Where:

- C = call option price
- N = CDF of the normal distribution
- S_t = spot price
- K = strike price
- r = risk-free interest rate
- t = time to maturity
- σ = volatility

When attempting to apply the Black-Scholes Formula to our simulation data, we discovered some inconsistencies between our product and the products Black-Scholes is meant to price. Notably, at time zero, when we sell this product, we will have no idea the value of the strike price or the spot price. The interest rate and volatility will also need to be estimated based on historical precedent.

In our attempt to reach a price from the Black-Scholes formula, we assumed the strike price was known. Then, we calculated the spot price to be the present value of the strike price. For the interest rate, we tried both the geometric average of our interest rates and the geometric average of the natural log of the interest rates. For volatility, we applied either the standard deviation or the geometric standard deviation. The results of the four combinations of these options are seen in the table below.

Figure 18: Table of Black-Scholes method prices using various inputs

	r = Geometric Average of Interest Rates	r = Geometric Average of the Natural Log of the Interest Rates
σ = Standard Deviation	\$2,607	\$11,577
σ = Geometric Standard Deviation	\$298	\$4,686

The price found using the geometric average of the interest rates and the geometric standard deviation is a reasonable price at \$298 but would leave us to go bankrupt a large percentage of the time. While the other prices would limit the instances in which bankruptcy occurs, they seem overly expensive and unlikely that a customer would be interested in purchasing this policy.

Because the strike price and the spot price are both calculated from the geometric average interest rates, we explored how changes in that geometric average impact the price. Next, we investigated the effects of changes in the volatility.

Figure 19: Table of Black-Scholes method prices using various inputs

at σ =.2	
r	Pt
5%	$4,836
7.5%	$2,518
10%	$1,145
15%	$150

at r=7.5%	
σ	Pt
0.1	$153
0.2	$2,518
0.3	$5,306
0.4	$7,520

There is an inverse relationship between changes in the geometric average and changes in the standard deviation. If you double the geometric average, you get nearly an identical price to when you halve the standard deviation. It is important to note that the price goes up as the interest rate goes down. This is important because the risk-free interest rate is a smaller number than the geometric average. The risk-free interest rate of the S&P500 averages around 4.5%, meaning the recommended price would be quite high, upwards of $5,000. Because of this, along with the fact that many of these values are unknown at the time of payment, we decided that using Black-Scholes is not an effective method to price this product.

Direct Simulation Method

As a complement to the prices found using the Black-Scholes formula, we used a method involving direct simulations to derive price values as well. For this method, we generated 50,000 independent simulations of our policy, each of which selected 30 random annual growth factors from the S&P500's history. This leads to all policies all having different growth factors, with a unique geometric average and strike price.

We used each policy's growth factors to calculate a premium that would cover the policy's payment, using the total S$&P500 yield. The total yield is adjusted to reflect any dividend payments that have been made. As previously mentioned, we are guaranteeing our customers the price yield of the S&P500. The long-term average difference between the total yield and the price yield is about 1.8%. Essentially, by only guaranteeing the price yield, we perform 1.8% better than our customers per year. This is actually a good thing for both us and our customers. This higher yield allows us to charge a smaller premium, making our product more affordable for our customers. If there is no payment, then we keep the extra profit, so that is good for us. Implementing this idea, we calculated the premium necessary to cover various percentiles of our 50,000 simulated policies.

Figure 20: Table of percentiles of direct simulation payment values

Percentile	Premium
100%	$ 29,448
99%	$ 6,707
95%	$ 4,166
90%	$ 3,145
80%	$ 1,992
50%	$ 0

Figure 20 displays the premium required to cover the given percentile ranges of loss. For coverage of 95th percentile losses, each policy would require a premium of $4,166. Considering that the annual investment is $1,000, this price is quite steep. It is very unlikely that a customer would be willing to purchase this product. Hopefully, there are other methods of acquiring capital to pay losses that will reduce the premium for our customers, and still generate a profit. Some potential solutions have come up in our work, but we have not had enough time to analyze them thoroughly. We will present these ideas in a later section about further research.

Figure 21: Table of percentiles of direct simulation payment values with dividend growth

Percentile	Payment
100%	$16,759
99%	$6,707
95%	$3,990
90%	$2,505
80%	$1,211
50%	$0

Comparing the direct simulation with the initial simulation, the direct simulation has a much wider range of variance. Since the volatility increases the results of the extreme outcomes, the payment distribution is further right skewed. Looking at Figure 21, the 90th percentile of the direct simulation payments is approximately $2,500. For the initial simulation payments displayed in Figure 22, a payment of approximately $2,500 is the 95th percentile instead. However, as you get further away from the extremes, the payment distributions converge. At 80% percentile, the two simulations have payments within $100 of each other, $1,211 and

$1,296, and the expected payments are near equal. The similarities between these two simulations give confidence that they are representative models of reality.

Figure 22: Table of percentiles from initial 10,000 simulation payment values

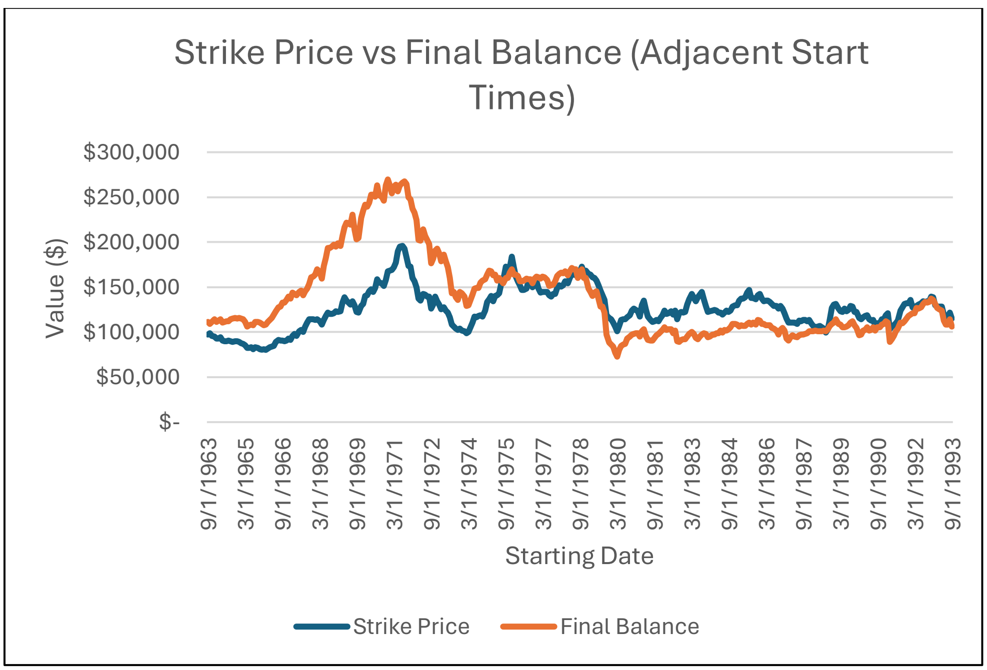
Percentile	Payment
100%	$4,333
99%	$3,206
95%	$2,457
90%	$1,974
80%	$1,296
50%	$0

Diversification Test

To determine if time alone would be a sufficient method of diversifying our product. We simulated 30-year policies to see how the variation in the effective date of each policy affected the totals. For the simulation, we selected the monthly interest values from a random point in the

S&P500's history. A new policy was set to start every month for the first 30 years, and each of these policy's annual interest rates would be compounded in 12-month intervals from when they started, meaning policies that started in the same year would have unique interest rates by combining months from different starting points.

Figure 23: Graph of strike price vs final account balance over time, sold at 1-month intervals



The result of this test strongly indicates that time is not a sufficient diversifier for this product. Roughly half of the policies finished with account balances above the strike price. However, all those policies were started from 1963 to 1978 in the simulation, while the values of policies starting from 1978 to 1993 were nearly all below their strike prices. Despite the fact that the monthly policy separation resulted in a unique set of interest rates, the order and size of those rates overlaps with the other policies. There are only two completely independent sets of interest rates: the first 30 years and the last 30 years. It is easy to imagine a scenario where both time extremes have general downwards trends, and a vast majority of policies are in the money. This small amount of variance for 60 years of policies is not sufficient diversification on its own for this policy to work. Some other form of diversification or hedging must be implemented.

Further Research

Introduction

Throughout the process of working on this project, we have come across many ideas that we believe would benefit this product greatly. We have explored some of these ideas, but unfortunately, we have fallen short of a solution. This section discusses our efforts with these ideas, in hopes of inspiring further research to improve the product.

Scoring Sequence Risk

One of our goals while conducting simulations was to determine a formula or method that gauges the magnitude of sequence risk based on the growth factors. In other words, we wanted to create a function where the input is a set of growth factors, and the output is the standard deviation of the account balances. The main idea was to use some property of the growth factors, like their standard deviation, to predict the account balance deviation through a linear formula, such as:

$$\sigma_{balance} = A \cdot \sigma_{growth\ factors} + B$$

To approximate the values of A & B in this function, we created 100 sets of 10 growth factors, using a random generation method that ensured a variety of standard deviations and geometric averages. For each of these 100 sets, we conducted 100 random-permutation simulations of an annuity policy to determine each sample standard deviation for the ending account balances. We measured the correlation between the input and output summary statistics to find which were most closely related, so we could use linear regression to generate a function.

Of the statistics we tested, the arithmetic standard deviation of the account balances was most closely correlated with the arithmetic standard deviation of the growth factors, with a coefficient of 0.718. The geometric standard deviation of the account balances was much clearer to predict, correlating best with the geometric standard deviation of the growth factors, with a coefficient of 0.990. Although this coefficient is more ideal, the output value of geometric standard deviation is a less applicable measurement of risk for account balances, since geometric statistics are meant for compounding values like growth factors, not account balances.

Regardless, we attempted to perform regression on the higher coefficient scoring method and ended with the following formula, but we recommend further research into this topic to find a better method.

$$\sigma_{balance} = 0.88 \cdot \sigma_{growth\ factors} + 0.12$$

Capital Investment

One complication that we discovered with our pricing methods is that the premium that our customers would have to pay for coverage is too high. With an expected payment of $8,477 per policy, a $1,000 premium only covers up to the 55th percentile of losses. This is the same as the customer making two investments at the beginning of the 30 years. To ensure that we have enough money to pay all losses, we would need $4,285 at the start of the policy. This is simply too much money to ask for and it would be very difficult to acquire customers.

One potential solution to this problem is to reach out to large capital investors. The investor would provide the remaining capital necessary to cover a given percentile of losses. This paid-in capital would be utilized to pay claims. If there are no losses, or the losses are covered by the customers' premiums, then the capital investor would be given all or the remaining portion of the total premiums. This way, the investor takes on all the risk, but gets all the reward. We would make money through various fees during the underwriting process.

Similarly to the customers, the price for the capital investors needs to be worthwhile for the coverage that they provide. In our example, we assume that the initial investment will grow at 9.4% annually. This includes the total dividend yield of the S&P500. If the capital investment were to be $1,000, this would cover up to the 90th percentile of loss. The table below displays the various potential capital investment amounts, percentile coverages, expected internal rate of returns (IRR), and best-case scenario internal rate of returns.

Figure 25: Table of investment amounts and percentile coverages

Investment Amount	Percentile Coverage	Expected IRR	IRR without Losses
500	83.64%	11.677%	13.481%
1000	90.65%	10.706%	11.957%
1500	95.56%	10.318%	11.279%
2000	98.36%	10.109%	10.889%
2500	99.60%	9.977%	10.634%
3000	99.97%	9.887%	10.454%
3500	100%	9.821%	10.320%

If the investor was to instead put their money in the S&P500 for the 30 years, their realized annual return would be 9.4%, the same as ours. Based on the graph above, the improvement that the investor would get is a minimal one to two percent. This is likely not enough upside for the investment to be worth taking the risk of covering the losses. Therefore, this is not the best option for getting enough money to cover the losses and another method must be found. This idea of searching for outside capital to cover losses would be excellent if it could be slightly more profitable and slightly less risky for an investor. Any adjustments that could be made to the approach that would produce this outcome would make for great further research for our project.

Policyholders' Dividend Payments

As stated earlier, we are guaranteeing the price yield to our customers. However, this does not mean that dividends will not be paid on their account. We are just choosing to ignore them in our calculations. So, what if there was a way to use these dividend payments as a supplement for premium? After a brief analysis, it seems that collecting all of these payments would provide coverage for most, if not all possible losses. In practice, this may be unreasonable. As our analysis indicates, the average final account balance for the simulation was over $100,000. This means that a dividend payment at the end of the 30 years, which is 1.8% on average, could be greater than $2,000. Keeping such a large part of the investment might not sit well with our customers. In a perfect world, we would retain as high a percentage of the dividend payments as possible, without deterring our customers.

If we were to implement this strategy, there would be one very important question to answer. Specifically, who is in possession of the account throughout the 30 years? If the answer is the customer, then they are essentially making a payment to us every year, and this is no different than a variable premium payment. This may not be received very well. If the answer is us, then we would enter a storm of regulations. Managing other people's money can be very risky, and as such, there are many rules to follow, making the entire process both more difficult and more stressful for us.

This is something that could be the difference in the marketability of our product, but it is important to consider every detail carefully. The ideas mentioned in this section could each provide enough work for an entirely new project, as there are seemingly endless possibilities for adding capital as a complement to premium.

Conclusion

We were able to observe and quantify the impact of sequence risk on a pre-picked set of 30 growth factors, and one could price various confidence intervals based on that estimation. In real world application this policy would not know what the growth factors will be at the time the policy is sold, making the strike price unknown and the product difficult to price. Since all policies sold at the same time receive the same random outcomes of the S&P500, the law of large numbers does not apply and creates difficulty with diversifying the product. This product could be more firmly established by future research into methods for hedging or alternative diversification. We were able to use methods of brute-force simulation and the Black-Scholes method to estimate price ranges for various levels of confidence for coverage, but a more exact value could be pinpointed once a hedging or diversification method is established to mitigate risk. Further efforts with a scoring method could also help towards securing specifics for pricing.

Appendices

Appendix A: Proof for Geometric Standard Deviation of Spread Increasing Method

To increase the spread of a set of values without changing the geometric average, the formula performed on every term of set A_0 is as follows:

$$a_{3i} = \left(\frac{a_{0i}}{Geo\left(A_0\right)}\right)^c \cdot Geo\left(A_0\right)$$

We'll state the before & after sets for the entire process as A_0 & A_3, written as follows:

$$A_0 = \left\{a_{01}, \ldots a_{0n}\right\}$$

$$A_3 = \left\{\left(\frac{a_{01}}{Geo(A)}\right)^c \cdot Geo(A), \ldots \left(\frac{a_{0n}}{Geo(A)}\right)^c \cdot Geo(A)\right\}$$

$$(Proven\ above):\ Geo\left(A_0\right) = Geo\left(A_3\right)$$

To track how this process affects the geometric standard deviation of a set, we'll start by writing out the definition for the geometric standard deviation of A_0:

$$GeoSD\left(A_0\right) = \exp\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\ln\left(\frac{a_{0i}}{Geo\left(A_0\right)}\right)^2\right)}\right)$$

Now we will state the same thing for set A_3, using its terms we defined above:

$$GeoSD\left(A_3\right) = \exp\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\ln\left(\frac{\left(\frac{a_{0i}}{Geo\left(A_0\right)}\right)^c \cdot Geo\left(A_0\right)}{Geo\left(A_3\right)}\right)^2\right)}\right)$$

Since Geo(A_1) = Geo(A_3), we can cancel them out in the larger fraction:

$$= \exp\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\ln\left(\left(\frac{a_{0i}}{Geo\left(A_0\right)}\right)^c\right)^2\right)}\right)$$

Now, this expression looks remarkably like the expression for GeoSD(A_0), except for the constant power c on the inside. However, since it is a constant, we can drive it to the outside of the equation through a few algebraic steps on the next page:

$$= \exp\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(c\left(\ln\left(\frac{a_{0i}}{Geo\left(A_0\right)}\right)\right)\right)^2}\right)$$

$$= \exp\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}c^2\left(\ln\left(\frac{a_{0i}}{Geo\left(A_0\right)}\right)^2\right)}\right)$$

$$= \exp\left(\sqrt{c^2 \cdot \frac{1}{n}\sum_{i=1}^{n}\left(\ln\left(\frac{a_{0i}}{Geo\left(A_0\right)}\right)^2\right)}\right)$$

$$= \exp\left(c \cdot \sqrt{\cdot \frac{1}{n}\sum_{i=1}^{n}\left(\ln\left(\frac{a_{0i}}{Geo\left(A_0\right)}\right)^2\right)}\right)$$

$$= \left(\exp\left(\sqrt{\cdot \frac{1}{n}\sum_{i=1}^{n}\left(\ln\left(\frac{a_{0i}}{Geo\left(A_0\right)}\right)^2\right)}\right)\right)^c$$

Now that c is on the outside, we can replace the entire inside with our original definition of GeoSD(A_0):

$$= GeoSD\left(A_0\right)^c$$

This establishes the geometric standard deviation of our new set to be equal to the old one raised to a power c; a pattern which we observed with the samples we tested this method with.

Appendix B: ChatGPT citation

To create our 10,000 random permutations in excel, we used a macro created by ChatGPT. The conversation in which this macro was generated is linked in the citation below:

OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model].

https://chat.openai.com/share/79731bec-4cd8-4184-b11c-b524c8ca0093

The macro generated:

```
Function GenerateRandomPermutation(rng As Range) As Variant

    Dim arr() As Variant

    Dim i As Long, j As Long

    Dim temp As Variant



    ' Read data into an array

    arr = rng.Value



    ' Randomize the array

    For i = 1 To UBound(arr, 2)

        j = Int((UBound(arr, 2) - i + 1) * Rnd + i)

        temp = arr(1, i)

        arr(1, i) = arr(1, j)

        arr(1, j) = temp

    Next i
```

```
    GenerateRandomPermutation = arr

End Function
```

Usage example:

```
=GenerateRandomPermutation($B$1:$AE$1)
```

This creates a 30-cell-long row of the values from B1 to AE1 in a random order.

References

Andrew Clare, James Seaton, Peter N. Smith & Stephen Thomas (2017) Reducing Sequence Risk Using Trend Following and the CAPE Ratio, Financial Analysts Journal, 73:4, 91-103, DOI: 10.2469/faj.v73.n4.5

S&P 500 Historical Prices by Month - Multpl

OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model].

https://chat.openai.com/share/79731bec-4cd8-4184-b11c-b524c8ca0093

(used to generate code for a macro)