Simulating Future Interest Rates? Judgment? Art? Science!

Zexing Hu, Yuting Tang, Ao Xu
Faculty Advisor: Hal Pedersen
Department of Statistics and Applied Probability - University of California, Santa Barbara

ABSTRACT

An Economic Scenario Generator (ESG) is employed to simulate data pertaining to the economy to help insurance regulators in establishing policy. In 2022, the National Association of Insurance Commissioners released a new ESG model (NAIC) and is considering replacing the previously used Academy's Interest Rate Generator (AIRG). To evaluate the two models, we first establish a set of acceptance criteria by employing standard econometric modeling techniques and researching key historical periods. Then, we compare and assess the performance of the two ESGs based on the criteria we developed, in order to discuss the strengths and weaknesses of the two models. Our findings indicate that the NAIC model produces more volatile results than the AIRG model.

INTRODUCTION

Our project aims to evaluate the performance of two ESG models, which are the AIRG model and NAIC model. Our initial step is to analyze historical data to gain insight into market behavior and form capital market assumptions. Based on the assumptions, we can establish acceptance criteria for comparing the two models in the later stages of the project.

The NAIC is an entity that provides expertise, data, and analysis for insurance commissioners to regulate the industry and protect consumers effectively. In 2017, the American Academy of Actuaries(AAA) stopped maintaining the ESG. Then, the NAIC Life and Annuity Task Force (LATF) authorized its vendor, Conning, to provide, maintain, and support the generator.

ESG is a computer-based model that generates economic scenarios. It provides several simulated instances of possible future values of various economic and financial variables, along with an analysis of the stochastic distribution of scenario outcomes. Simulated results produced by the ESG can help insurance regulators understand the risks and uncertainties that exist within the economy, which can then be used to inform policy decisions.

There are three component models in an ESG, which are: 1)Treasury Interest Rate Model; 2) Bond Fund Model; and 3) Equity Return Model. Among the three components, the Treasury Interest Rate Model is the core model of the generator. It is a standard finance Cox-Ingersoll-Ross 3-factor interest rate model that is used to simulate future treasury yield rates. The Bond Fund Model is still in the development process. And the Equity Model is used to provide returns for representative equity funds offered in US Life Insurance products.

DATA

I. Stocks, Bonds, Bills, and Inflation (SBBI)

Length	From 1926.01 to 2022.12	
Variables	Large Cap (S&P 500) TR	Long Term Govt Bond TR (20yr)
	Intermediate Govt Bond TR (5yr)	Long Term Corp Bond TR (20yr)
	US Treasury Bill (1 Month) TR	US CPI (Inflation)
	Long Term Govt Yield (20yr)	Intermediate Govt Bond Yield (5yr)

Table 1: SBBI Data Description

II. Federal Reserve Economic (FRED)

Length	From 1954.01 to 2023.01
Variable	US 3 Month Treasury Yield

Table 2: FRED Data Description

Full name: 3-Month Treasury Bill Secondary Market Rate, Discount Basis, Percent, Monthly, Not Seasonally Adjusted

III. Academy Interest Rate Generator (AIRG)

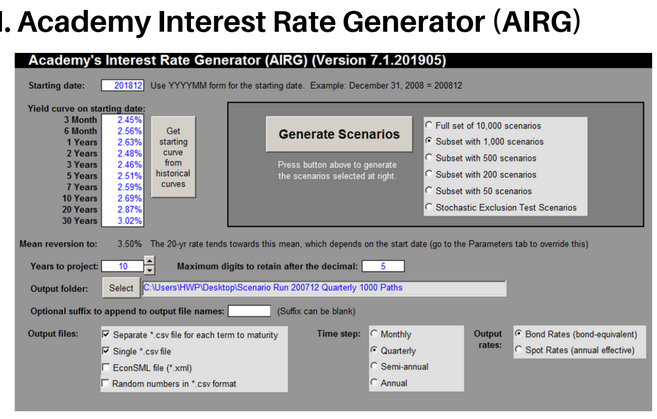

Figure 1: AIRG Simulator User Page

IV. National Association of Insurance Commissioners Economic Scenario Generator (NAIC)

Since the Conning company did not make their model public, we decide to use the test data (NAIC ESG Field Test #1a: Scenario Sets and Statistics: 1000 Path Subset 1a Conning GFF Baseline Equity 123121) released by the company to conduct our analysis.

METHODS

I. Treasury Interest Rate Component

There are two approaches used to analyze the rate level produced by the simulators. The first method involves comparing the Percentiles Exponentially Weighted (PEWs) of historical and simulated data, while the second involves using the backtesting method.

PEWs:

The underlying idea is to give greater importance to the most recent data and less importance to older data when comparing the distribution of data.

	Monthly Data Point	Annual Data Point	Simulation Path	Dimension
SBBI (1926-2022)	1164	97	NA	1 X 1164
FRED (1954-2023)	829	69	NA	1 X 829
AIRG (1954-2022)	828	69	1000	1000 X 828
NAIC (Field Test 1a)	1200	100	1000	1000 X 1200

Table 3: SBBI and FRED refer to the historical data sources. As the NAIC model has not been made public, we utilized the test file released by the operating company.

As our datasets have varying dimensions, the weight calculation process differs slightly. The procedure is outlined below:

For Historical Data:

$$w_i = (1-\alpha)^{n*12-i}$$

$$nw_i = \frac{w_i}{\sum_i^{n*12} w_i}$$

$$cbind(data, \bar{NW})$$

For Simulated Data:

$$W[1,] = (w_1, w_2, \ldots w_N)$$
$$rep(W[1,]) \times 1000$$
$$NW[i,j] = \frac{W[i,j]}{\sum W}$$
$$NW \sim (1000 \times N)Matrix; \sum NW = 1$$
$$NAIC \sim (1000 \times 1200)Matrix; AIRG \sim (1000 \times 828)Matrix;$$
$$cbind(unlist(NAIC), unlist(NW))$$

The alpha value is calculated by setting a half-life of 48.5 years, resulting in an alpha value of approximately 0.002. Once the normalization weights have been appended to the original dataset following the aforementioned steps, we sort the original data and compute the cumulative sum of the normalized weights. Subsequently, we employ a *which* statement to determine the desired quantile.

Backtesting:

Backtesting, utilizing historical data, is the evaluation process we used to assess the reliability and validity of the two predictive models. This method involved a thorough comparison of simulated data against historical data for US 3-month Treasury Bills and US 20-year Treasury Bills at the reference year of 2022. We first calculate the minimum and maximum values, as well as various percentiles (10%, 20%, 50%, 80%, and 90%), for simulated data produced by the two models. We then compare these values to the corresponding percentiles in the historical data. Through careful examination of how well the range of simulated paths aligns with the historical data, we can assess the effectiveness of the two ESG models in accurately representing the historical data.

II. Equity Return Component

The statistic of Equity Risk Premium is calculated to compare the performance of the two model. The formula is defined below:

$$Equity\ Risk\ Premium = \frac{Large\ Cap\ TR}{US\ Treasury\ TR_{1mo}} - 1$$

Note: The AIRG model does not simulate 1-mo US Treasury TR. Thus, we use 3-mo US Treasury TR to calculate the equity risk premium for the AIRG model.

III. Bond Fund Component

The statistic of Bond Risk Premium is calculated to compare the performance of the two model. The formula is defined below:

$$Bond\ Risk\ Premium = \frac{Bond\ Return}{US\ Treasury\ TR_{1mo.}} - 1$$

Note: The AIRG model does not simulate 1-mo US Treasury TR. Thus, we use 3-mo US Treasury TR instead. And since the AIRG model does not simulate long-term government bond return, we are not able to calculate the bond risk premium for long-term bond for the model.

RESULTS

I. PEWs

US Treasury Yield - 3 Mo. (PEWs)			
Percentile	Historical Data	AIRG	NAIC
Min	-0.01%	0.01%	-1.15%
1%	0.01%	0.01%	-0.64%
5%	0.04%	0.39%	-0.35%
15%	0.14%	0.82%	-0.03%
30%	1.56%	1.23%	0.30%
50%	3.55%	1.73%	1.94%
70%	5.11%	2.34%	4.37%
85%	6.97%	3.10%	7.13%
95%	9.18%	4.33%	11.02%
99%	14.22%	6.34%	16.33%
Max	15.52%	16.78%	38.09%

Table 4: Pews Result for 20 Year US Treasury Yield

US Treasury Yield - 20 Yr (PEWs)			
Percentile	Historical Data	AIRG	NAIC
Min	0.97%	0.92%	0.14%
1%	1.17%	1.44%	0.38%
5%	2.00%	1.81%	1.00%
15%	2.46%	2.18%	1.81%
30%	3.03%	2.54%	2.72%
50%	4.38%	2.95%	3.99%
70%	6.23%	3.45%	5.60%
85%	8.10%	4.10%	7.55%
95%	10.95%	5.16%	10.45%
99%	13.48%	6.88%	14.79%
Max	14.82%	16.86%	33.52%

Table 5: Pews Result for 3 Month US Treasury Yield

Upon examining the overall figures, we discover that both models possess the ability to replicate low rates and simulate extreme value. These two outcomes are in agreement with the historical market behavior.

II. Backtesting

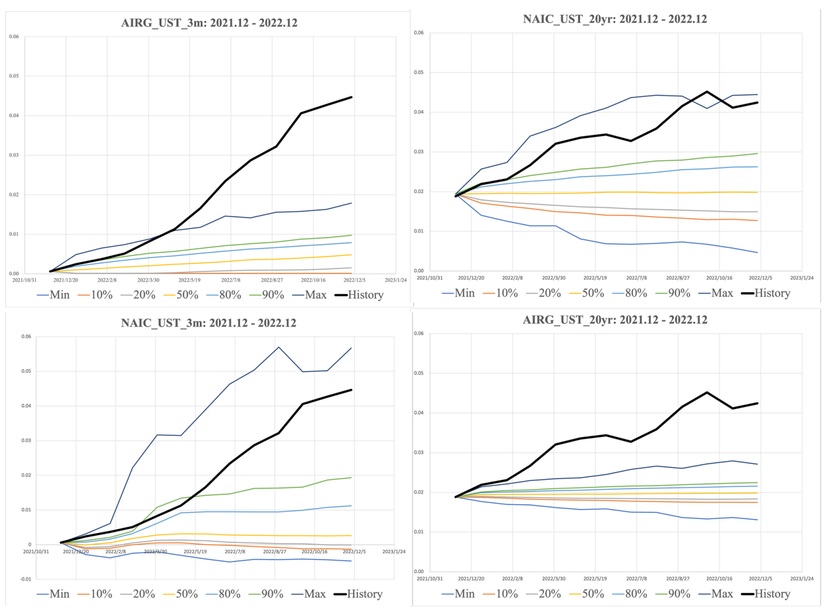

Figure 2: Backtesting Results for 3 Month US Treasury Yields (Left) and 20 Year US Treasury Yields (Right)

Throughout the year 2022, the recorded data for the US 3-month Treasury Yields and US 20-year Treasury Yields demonstrated a significant increase. As a result, these figures were beyond the expected range for the majority of the simulated paths in the AIRG model. However, the NAIC model is able to generate results that reflect the fluctuation of the historical data.

III. Equity Risk Premium

Equity Risk Premium (Annual)			
Percentile	Historical Data	AIRG	NAIC
Min	-44%	-61%	-62%
1%	-38%	-30%	-29%
5%	-24%	-19%	-18%
10%	-14%	-13%	-12%
25%	-3%	-4%	-3%
50%	9%	7%	7%
75%	22%	16%	15%
90%	31%	26%	25%
95%	37%	33%	33%
99%	51%	48%	51%
Max	54%	160%	156%

Table 6: Equity Risk Premium Result

From examining the table, we can deduce that the distribution of simulated data from both models within the 1st and 99th percentile is very similar to that of the historical data. However, the tail values of the simulated data differ greatly from the actual data.

IV. Bond Risk Premium

Bond Risk Premium - Intermediate (5YR) (Annual)			
Percentile	Historical Data	AIRG	NAIC
Min	-9%	-26%	-31%
1%	-7%	-8%	-16%
5%	-5%	-5%	-11%
10%	-3%	-4%	-8%
25%	-2%	-2%	-3%
50%	2%	1%	1%
75%	4%	3%	6%
90%	9%	5%	11%
95%	10%	7%	15%
99%	12%	10%	22%
Max	17%	29%	49%

Table 7: Bond Risk Premium Result (Intermediate Government Bond)

Bond Risk Premium - Long-Term (20YR) (Annual)			
Percentile	Historical Data	AIRG	NAIC
Min	-15%	NA	-41%
1%	-14%	NA	-24%
5%	-11%	NA	-17%
10%	-8%	NA	-13%
25%	-3%	NA	-6%
50%	3%	NA	2%
75%	7%	NA	9%
90%	15%	NA	17%
95%	22%	NA	23%
99%	27%	NA	34%
Max	27%	NA	76%

Table 8: Bond Risk Premium Result (Long-term Government Bond)

The distribution of simulated data from both models within the 5th and 95th percentile is very similar to that of the historical data, but the tail values of the simulated data differ greatly from the actual data.

DISCUSSION

Based on the analysis of the exponentially weighted percentiles, we discover that both simulators can replicate low rates and simulate extreme values, which is consistent with historical market behavior. However, when we look at the tail distribution of the simulated data from both models, we find that the rates produced by the NAIC model are more unstable compared to the AIRG Model.

Based on the outcomes of the backtesting, the NAIC model is capable of predicting a broader range of data that encompasses the range of historical data. On the other hand, the AIRG model generates more stable data but does not produce data that accurately reflect certain unusual market behavior.

Based on the results from equity/bond risk premium, we discover that the majority of the simulated data from both models match the distribution of the historical data. Nevertheless, the minimum and maximum values produced by both models suggest that the distribution of the simulated data has a heavy tail on both sides. This implies that the models may occasionally generate outliers that users should be cautious of. If users intend to utilize the simulated data to make decisions, they should consider disregarding those data points.

In conclusion, we propose that both models have their advantages and disadvantages, and it is difficult to determine which one is better. However, we consider the AIRG Model a better choice for users who want more conservative results compared to the NAIC model.

REFERENCES

[1] *NAIC RISK-BASED CAPITAL.* (2022, December 6). Retrieved from https://content.naic.org/cipr-topics/risk-based-capital#:~:text=The%20RBC%20requirement%20is%20a,in%20proportion%20to%20its%20risk

[2] Pasaogullari, M., & Meyer, B. (2010). *Simple Ways to Forecast Inflation: What Works Best?* FEDERAL RESERVE BANK of CLEVELAND. Retrieved from https://doi.org/10.26509/frbc-ec-201017

[3] *Why does the Federal Reserve aim for inflation of 2 percent over the longer run?* BOARD OF GOVERNORS of the FEDERAL RESERVE SYSTEM. (2020). Retrieved from https://www.federalreserve.gov/faqs/economy_14400.htm

[4] Boesler, M., & Bloomberg. (2022, September 21). *The Fed raises interest rates by 75 basis points for the third time this year.* FORTUNE. Retrieved from https://fortune.com/2022/09/21/fed-raises-interest-rate-75-basis-points/

[5] Kehrberg, J., & Richardson, L. (2022). (rep.). *Interest Rates - Stylized Facts and Acceptance Criteria* (pp. 1–43). National Association of Insurance Commissioners (NAIC) Life Actuarial (A) Task Force (LATF).

[6] Duff & Phelps. (2016). *2016 SBBI Stocks, Bonds, Bills, and Inflation.* John Wiley & Sons, Inc., Hoboken, New Jersey.

[7] *Scenario Files.* National Association of Insurance Commissioners. (2023). Retrieved from https://naic.conning.com/scenariofiles

[8] *Economic Scenario Generators.* AMERICAN ACADEMY of ACTUARIES. (2023). Retrieved from https://www.actuary.org/content/economic-scenario-generators

ACKNOWLEDGEMENTS

We express our gratitude to Professor Hal Pedersen for his outstanding mentorship throughout the entirety of this research endeavor. Additionally, we acknowledge the support provided by the University of California, Santa Barbara.