# MS-E2177 — Seminar on Case Studies in Operations Research
# Evaluating Interest Rate and Financing Risk of Sovereign Debt
# Final Report

Mika Viirret (Project Manager), Kerkko Konola, Kaarle Mäkelä, Topias Malinen, Elmer Heino

June 11, 2026

# Contents

# 1 Introduction

## 1.1 Background

The Finnish State Treasury manages the central government debt portfolio with the objective of ensuring continuous access to funding and meeting all obligations at the lowest feasible cost over time, subject to acceptable risk. Debt management decisions are forward-looking: they shape the maturity profile, refinancing schedule, and interest cost exposure of the government for many years. Because future financing conditions are uncertain, debt strategy evaluation is fundamentally a stochastic decision problem, where policy choices should be assessed under a range of plausible macro-financial scenarios rather than through point forecasts alone (Bolder 2003; Belton et al. 2018).

The current operating environment increases the importance of a robust risk framework. Interest rates have exhibited large and rapid movements across the yield curve, and macroeconomic conditions have been volatile. These dynamics affect both near-term debt service costs and the longer-term sustainability of the debt path. From a sovereign risk-management perspective, it is not sufficient to focus on expected outcomes: tail events and adverse scenarios can dominate welfare-relevant losses and constrain fiscal space. Modern approaches therefore emphasize distributional risk measures, stress testing, and strategy design that explicitly accounts for tail risk and sustainability constraints (Zenios et al. 2021).

In this setting, interest rate risk must be understood broadly. It includes (i) near-term increases in debt servicing costs due to rate rises and refinancing at higher yields, (ii) persistent changes in the average cost of funding that affect medium-term budget balances, and (iii) sustainability risks that arise when the effective interest rate on the debt stock exceeds the growth rate of the economy over extended periods. The last component is closely related to the interest–growth differential: when the effective interest rate persistently exceeds growth, debt dynamics can deteriorate even under moderate primary imbalances, increasing the relevance of prudent maturity management and risk controls (Zenios et al. 2021; Belton et al. 2018).

From a decision-making perspective, the State Treasury must evaluate alternative issuance and refinancing strategies (for example, changing issuance across maturities, smoothing refinancing needs, and setting limits for interest expenditure risk) and assess whether strategies remain robust under uncertainty (Bolder 2003; Belton et al. 2018). This project addresses this need by developing a simulation-based, decision-relevant framework that generates coherent scenarios for the term structure of interest rates and key macro drivers, maps those scenarios into debt cash flows and debt stock dynamics, and produces interpretable risk metrics for strategy comparison.

Methodologically, the framework combines stochastic scenario generation with structured evaluation of financing cost and risk. The stochastic calculus and continuous-time modelling toolkit provides the theoretical foundation for simulating risk-factor dynamics and connecting them to pricing and cash-flow processes (Shreve 2004). For interest-rate dynamics, term-structure modelling provides a principled way to generate yield-curve scenarios consistent with no-arbitrage and empirically relevant co-movements (Brigo and Mercurio 2006). Finally, the sovereign debt literature demonstrates how scenario-based analysis and optimization can be used to manage tail risk and incorporate sustainability considerations directly into debt-financing decisions (Zenios et al. 2021).

### 1.2 Objectives

The primary objective of this project is to develop a simulation-based, decision-relevant framework for evaluating sovereign debt portfolios under macro-financial uncertainty. The framework is designed to support the State Treasury in comparing alternative debt management strategies by quantifying their implications for expected financing costs, refinancing needs, and downside risk, with emphasis on transparent assumptions and interpretable outputs (Bolder 2003; Belton et al. 2018).

More specifically, sub-objectives of the project are on follows:

- **Interest rate risk assessment.** Generate coherent yield-curve scenarios and quantify how changes in the term structure translate into distributions of debt servicing costs across horizons (Brigo and Mercurio 2006).
- **Refinancing and funding risk assessment.** Model gross financing needs and refinancing profiles under alternative issuance rules, and quantify the probability and severity of adverse refinancing outcomes (Bolder 2003).
- **Debt sustainability analysis.** Evaluate debt-to-GDP dynamics under joint interest-rate and macroeconomic scenarios, and assess tail-risk and sustainability constraints in a distributional setting (Zenios et al. 2021).
- **Strategy comparison under constraints.** Compare issuance and maturity-structure strategies using cost–risk trade-offs (for example, expected cost versus tail-risk metrics) while respecting feasibility constraints relevant for sovereign issuance (Belton et al. 2018).
- **Model transparency and interpretability.** Provide decision-ready outputs, including fan charts, threshold breach probabilities, and clearly documented assumptions, so that results can be scrutinized and used in policy discussion (Zenios et al. 2021; Bolder 2003).

## 2 Literature review

### 2.1 Sovereign debt risk management and sustainability

The sovereign debt strategy problem has been extensively studies before. The problem on how to optimize the term maturity structure and strategy concerns all indebted governments (Brigo and Mercurio 2006; Hamilton 1989; Kim and Nelson 1999).

One influential model on this topic, is the one made about US sovereign debt by the The Hutchins Center on Fiscal and Monetary Policy. In the cited model they have a macroeconomic model, which is "off the shelf" meaning it uses established practices: a regression model with some stochastic variables. This macro model influences the interest rate model, which is a simple regression model known in literature as the Adrian, Crump, and Moench model. The impact of treasury issuance strategy on term premia is neglected. In the fiscal block, they apply another simple regression equation which depends on macroeconomic variables. This equation is adjusted to reflect on an established forecast of government debt-to-GDP ratio from an external provider. In addition the model also includes an optimization package, where the debt issuance strategy is modeled as linear combination of kernels, each kernel reflection some specified maturity structure. (Belton et al. 2018)

Another comprehensive study on this topic is the working paper by the Canadian government on their sovereign debt. In the paper they have thoroughly considered the dynamics of sovereign debt and considered various stochastic variables, such as interest rates, government expenditure and financing requirement, debt strategy and maturity strategy, including correlations and uncertainties involved in these variables. (Bolder 2003)

Europe has indebted governments, for whom the topic is highly relevant. Zenios et al. (2021) considered the problem in the European context. They use a similar framework as other, modeling interest rates, state financing requirements and optimizing the debt strategy. It incorporates a risk measure, a rich set of instruments and stock and flow dynamics. It optimizes dynamic financing strategies also incorporating the existing debt stock. Gross financing needs determine debt dynamics and instruments are used to fulfill it. Uncertainty is represented through a discrete time- and state-space scenario tree. Interest rates are modeled as the sum of a risk-free component and endogenous risk and term premia that depend on the debt ratio and the maturity of the issued instrument.

Our model takes an approach inspired by these models but with unique characteristics.

### 2.2 Interest rate scenario generation

To build a model for optimizing the the portfolio and maturity structure of debt, the model needs to capture interest rates and their evolution over time. Many models exists for to achieve this. Some common families of models are Gaussian short rate models (the Vasicek family), the more broad HJM (Heath–Jarrow–Morton) framework and Hull-White model, and finally the empirical factor models such as VAR where PCA can be used for dimensionality reduction. (Brigo and Mercurio 2006)

The Vasicek family belongs to the short rate family of models, the common one factor and two factor Vasicek models are based on defining the dynamics of a short rate using one or more stochastic OU (Ornstein-Uhlenbeck) processes. The advantage of such short rate models is that it is possible to derive closed form formulas for bond prices.

Short rate models have both advantages and disadvantages. There are also models, that are based on stochastic processes, but that are not based on modeling short rate dynamics but instead for example the forward rate as in the HJM framework as

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t). \tag{1}$$

The HJM framework is more general in that almost any interest rate model can be derived from it. (Brigo and Mercurio 2006)

For risk management and modeling purposes, the main advantages of this class of models is that they can be parametrized for various purposes, such as for modeling the possibility of negative rates, which can indeed occur in the context of sovereign debt. Parameter searches can be performed on the parameter space to generate scenarios allowing for the decision maker to incorporate his perspective in the model.

Another class of models consists of more general statistical models such as vector autoregression. They are common in economics and finance, and have been used in previous analyses of sovereign debt (see, e.g., Belton et al. 2018). However, a main challenge is that they are fitted to historical data, so a more creative parametrization of the decision maker's point of view can become rather challenging.

### 2.3 Macroeconomic forecasting and debt dynamics

When the state's spending exceeds revenue, the state must take on debt. The key concern is the debt-to-GDP ratio, and whether it stays stable over time. Depending on the effective interest on debt, and the growth rate of the economy, the state can run either a surplus or a deficit and the debt-to-GDP ratio will remain the same. This dynamic is governed by the debt dynamics equation:

$$\frac{d}{dt}\left(\frac{\text{Debt}}{\text{GDP}}\right) = (r-g)\frac{\text{Debt}}{\text{GDP}} + \frac{\text{Primary Deficit}}{\text{GDP}},$$

where $r$ is the effective interest rate and $g$ is nominal GDP growth. Abbas, Pienkowski, and Rogoff 2019

The economy changes between growth periods when the GDP grows and recessionary ones when is does not grow or even contracts. These are called business cycles, which present unique needs for macroeconomic forecasting because simple linear autoregressive models might not capture complicated dynamics. Markov regime-switching models, which allow growth rates to transition between discrete states with some specified transition probabilities, provide a solution to such challenges. (Hamilton 1989; Kim and Nelson 1999).

Bayes autoregressive models with plenty of variables, like the one at the Bank of Finland, have also been successfully used to forecast GDP with good results. In Bayes autoregressive models the variables of the model are treated as random variables that follow some prior distribution, while in regular autoregression these variables are simple fixed values. Their main advantage lies in their ability to handle limited amount of information while incorporating a large number of variables. (Itkonen and Juvonen 2017).

In macroeconomic literature the relationship between rates and GDP growth have been studied (see, e.g., Lee and Werner 2018).

### 2.4 Risk metrics, decision criteria and strategy

The economic literature provides a rather limited amount of insight when it comes to optimizing sovereign debt. (Abbas, Pienkowski, and Rogoff 2019). Limitations in literature make sovereign debt management a challenge. However, some precedents indeed do exits.

Sovereign debt strategy evaluation requires decision criteria that measure both typical performance and adverse outcomes. Common decision variables include the maturity structure of issuance and constraints that control issuance volatility and concentration. Outcomes of interest include interest expenditure, refinancing needs, and sustainability metrics such as Debt/GDP and GFN/GDP (International Monetary Fund and World Bank 2014; Escolano 2010; Missale 1999).

A central methodological point is that expected-cost minimisation alone may favour strategies that increase refinancing exposure. Risk-aware frameworks therefore incorporate tail-risk measures and stress testing. A coherent tail-risk measure is Conditional Value-at-Risk (CVaR), which admits a convex optimization representation and is widely used in stochastic optimization (Rockafellar, Uryasev, et al. 2000). In sovereign applications, CVaR-type objectives or constraints can be applied to interest expenditure, refinancing need measures, or sustainability losses, and the confidence level reflects policy risk tolerance (Zenios et al. 2021). Scenario-tree formulations are common in multistage stochastic programming, but scenario construction is itself non-trivial; alternative Monte Carlo style scenario sets require convergence and stability checks to ensure that tail estimates are numerically robust (Høyland and Wallace 2001).

# 3 Data & methods

## 3.1 Data preparation for growth model

The GDP model uses datasets provided by client including Finnish bond yields with 2 year, 5 year, and 10 year maturities, GDP deflator, and nominal GDP data per capita with quarterly resolutions. Additionally, GDP model uses G2++ model yield curve scenarios (2 year, 5 year, and 10 year) with monthly resolution.

**Frequency alignment and granularity.** The model makes quarterly frequency forecasts, which means that the G2++ yield curve forecasts need to be resampled to quarterly resolution. Resampling is performed by taking the mean of the monthly values.

**Missing data handling.** Missing data is handled using the following complementary strategies.

First, the datasets (GDP, deflator, and interest rates) are merged by an inner join using the `Date` column. This ensures that only complete and overlapping observations are used for model training, which reduces the risk of biased estimates.

Second, the Markov regime forecaster log-differences the time-dependent training data to convert the historical series into stationary form. This transformation produces a `NaN` value for the first row of the historical dataset, which is dropped before hyperparameter optimization starts.

**Sanity checks and validation.** Overlapping verification was carried out by checking the lengths of the differenced datasets and confirming that they matched. The lagged interest-rate structure was aligned by two periods relative to each GDP observation at time $t$. This is important for the Markov regime specification, in which state transition probabilities depend on correctly aligned historical state information.

## 3.2 Model requirements

To evaluate sovereign debt strategies over horizons of up to 50 years, the model must satisfy several requirements:

1. **Controllable long-run level.** The decision-maker must be able to set the average short-rate level to reflect a policy view or a baseline macro assumption (e.g. a long-run 2-year yield of 3%).
2. **Term premium (yield-curve slope).** The model must produce an upward-sloping yield curve at $t$=0 whose steepness is independently controllable. We choose the 10y–2y spread as a key input: it determines the cost–risk trade-off between short-term and long-term issuance. It is essential that the term premium does not arise merely from upward-trending rate expectations; instead, there should be a built-in risk premium. While this breaks the abitrage-free mechanics, it is essential for generating reasonable yield curves for the simulation.
3. **Independent control of curve slope and rate trend.** A positive yield-curve slope must not force the short rate to drift upward over the simulation horizon. This distinction is critical:

there is a need for scenarios where the curve is steep today, with rates remain stationary in the long run, as well as scenarios in which rates trend upward or downward independently of initial steepness.

4. **Cross-maturity correlation structure.** The model must capture the empirical observation that yields at different maturities are positively correlated but not perfectly so: i.e., while short-rate shocks decay along the curve, persistent shocks affect all maturities. A two-factor structure is the minimum needed to reproduce this (Brigo and Mercurio 2006).
5. **Volatility and mean reversion.** Factor volatilities govern the dispersion of interest-cost outcomes, and the speed of mean-reversion controls how long rate shocks persist. Both must be parameterizable so that the decision-maker can examine tail-risk scenarios at different confidence levels.
6. **Analytical tractability.** Bond prices must be available in closed form (affine in the state variables) so that yield curves can be computed without nested Monte Carlo. This is essential for computational feasibility in the evaluation of thousands of paths over long horizons.

The G2++ model, a two-factor Gaussian short-rate model introduced by (Brigo and Mercurio (2006)), satisfies requirements 1, 2, and 4–6 in its standard (arbitrage-free) formulation. In the standard G2++, the short rate follows $r_t = x_t + y_t + \varphi(t)$, where $x_t$ and $y_t$ are correlated Ornstein–Uhlenbeck processes and $\varphi(t)$ is a deterministic function calibrated to exactly match the observed initial term structure. The affine structure yields closed-form zero-coupon bond prices $P(t,T) = \exp\big(A(t,\tau) - B_x(\tau)\,x_t - B_y(\tau)\,y_t\big)$, where $\tau = T - t$ (Brigo and Mercurio 2006).

However, in the standard arbitrage-free G2++ model, a single parameter simultaneously controls both the yield-curve slope and the calendar-time trend of the short rate. A positive initial spread forces $\varphi'(t) > 0$, causing the expected short rate to rise indefinitely, which is unrealistic for 50-year debt-management scenarios and violates requirement 3.

To address this, we extend the G2++ model with a *decoupled* deterministic structure that replaces the single shift function $\varphi(t)$ with three independently controllable components:

- $\ell_0$ (`level0`) — the baseline short-rate level at $t$=0,
- $\gamma$ (`time_trend`) — calendar-time drift in the short rate,
- $\kappa$ (`curve_slope0`) — a yield-curve slope adjustment that enters bond pricing but has no counterpart in the short rate.

The short rate becomes $r_t = x_t + y_t + \ell_0 + \gamma\,t$, where $\gamma$ can be set to zero for a stationary long-run level. The slope parameter $\kappa$ enters only through the bond-pricing function $A(t,\tau)$, adding a term $-\frac{1}{2}\,\kappa\,\tau^2$ that produces a linear slope in yields versus maturity. This allows an upward-sloping curve without any implied rate trend.

However, this decoupling comes at the cost of internal consistency: because $\kappa$ affects bond prices but not the short rate, the model is *not* arbitrage-free. Short-rate paths and bond prices are not fully mutually consistent in the sense of classical no-arbitrage pricing theory. This is an intentional and documented trade-off. For derivative pricing, in which arbitrage-free consistency is essential, the standard G2++ model formulation should be used. For scenario-based debt-management analysis (CFaR, CDeaR), where the purpose is to generate plausible and controllable yield-curve scenarios rather than prices. For contingent claims, this trade-off is acceptable and the additional parametric flexibility is valuable (Brigo and Mercurio 2006; Hamilton 1989; Kim and Nelson 1999).

### 3.3 Baseline for model specification

The stochastic interest rate engine is based on the G2++ two-factor Gaussian short-rate model (Brigo and Mercurio 2006), extended with the decoupled deterministic structure described in Section 3.3.1.

**Factor dynamics.** Two latent factors $x_t$ and $y_t$ follow correlated Ornstein–Uhlenbeck processes with zero long-run mean:

$$dx_t = -a_x\, x_t\, dt + \sigma_x\, dW_1(t), \tag{2}$$

$$dy_t = -a_y\, y_t\, dt + \sigma_y\, dW_2(t), \tag{3}$$

where $\langle dW_1,\, dW_2 \rangle = \rho\, dt$. The deterministic component governs the level of the yield curve, while the stochastic factors capture mean-reverting deviations around that level at different persistence scales.

**Short rate.** The instantaneous short rate is

$$r_t \;=\; x_t + y_t + \ell_0 + \gamma\, t, \tag{4}$$

where $\ell_0$ sets the baseline level at $t{=}0$ and $\gamma$ is the calendar-time trend. By construction, the slope parameter $\kappa$ does not appear in the short rate.

**Bond pricing.** The price at time $t$ of a zero-coupon bond maturing at $t + \tau$ is

$$P(t,\, t{+}\tau) \;=\; \exp\!\big(\, A(t,\tau) - B_x(\tau)\, x_t - B_y(\tau)\, y_t \,\big), \tag{5}$$

with loading functions

$$B_x(\tau) = \frac{1 - e^{-a_x \tau}}{a_x}, \qquad B_y(\tau) = \frac{1 - e^{-a_y \tau}}{a_y}. \tag{6}$$

These are identical to the standard G2++ model specification (Brigo and Mercurio 2006). The log-price intercept is

$$A(t,\tau) \;=\; -\big[\, \ell_0\, \tau + \gamma\, t\, \tau + \tfrac{1}{2}(\gamma + \kappa)\, \tau^2 \,\big] \;+\; \tfrac{1}{2}\, V(\tau), \tag{7}$$

where

$$V(\tau) = \mathrm{Var}\bigg( \int_0^{\tau} (x_s + y_s)\, ds \bigg) \tag{8}$$

is the variance of the integrated stochastic factors, available in closed form from the standard G2++ formulae (Brigo and Mercurio 2006).

The first bracket is the deterministic contribution to bond pricing. The term $\gamma\, t\, \tau$ makes bond prices depend on calendar time, reflecting the accumulated effect of the short-rate trend. The term $\frac{1}{2}(\gamma + \kappa)\tau^2$ determines the deterministic linear slope component of the yield curve as a function of maturity. Here $\gamma$ contributes the part implied by the short-rate trend, while $\kappa$ acts as an additional slope adjustment that enters bond pricing only. Because $\kappa$ has no counterpart in (4), the model loses the full internal consistency of the arbitrage-free G2++ formulation, as discussed in Section 3.3.1.

**Yields.** Continuously compounded yields follow directly:

$$y(t,\tau) \;=\; -\frac{\ln P(t,t{+}\tau)}{\tau} \;=\; \ell_0 + \gamma\, t + \tfrac{1}{2}(\gamma+\kappa)\,\tau + \frac{B_x(\tau)}{\tau}\, x_t + \frac{B_y(\tau)}{\tau}\, y_t - \frac{V(\tau)}{2\tau}. \tag{9}$$

At $t{=}0$ with $x_0 = y_0 = 0$, the initial yield curve simplifies to

$$y(0,\tau) \;=\; \ell_0 + \tfrac{1}{2}(\gamma+\kappa)\,\tau - \frac{V(\tau)}{2\tau}. \tag{10}$$

Thus, the deterministic part of the initial curve is linear in maturity, while the variance correction introduces additional non-linear maturity dependence.

**Parameter interpretation.** Table 1 summarises the model parameters and their roles.

**Table 1:** G2++ decoupled model parameters.

| Parameter | Symbol | Role |
|---|---|---|
| Mean-reversion speeds | $a_x,\ a_y$ | Rate at which shocks decay; $a_x \gg a_y$ gives a fast and a slow factor |
| Factor volatilities | $\sigma_x,\ \sigma_y$ | Dispersion of short-rate and yield scenarios |
| Factor correlation | $\rho$ | Cross-factor co-movement; affects cross-maturity correlations |
| Baseline level | $\ell_0$ | Baseline short-rate level at $t{=}0$ |
| Calendar-time trend | $\gamma$ | Drift in the short rate; set to 0 for a stationary long-run level |
| Curve slope adjustment | $\kappa$ | Additional yield-curve slope component, independently controllable relative to $\gamma$ (bond pricing only) |

## 3.4 Simulation and calibration

**Exact discretization.** The OU processes (2)–(3) are simulated using exact discretization. Over a time step $\Delta t$, the conditional distributions are Gaussian:

$$x_{t+\Delta t} = x_t\, e^{-a_x \Delta t} + \sigma_x \sqrt{\frac{1-e^{-2a_x\Delta t}}{2a_x}}\;\varepsilon_1, \tag{11}$$

$$y_{t+\Delta t} = y_t\, e^{-a_y \Delta t} + \sigma_y \sqrt{\frac{1-e^{-2a_y\Delta t}}{2a_y}}\;\varepsilon_2, \tag{12}$$

where $(\varepsilon_1, \varepsilon_2)$ are standard normal with correlation $\rho$, obtained via a Cholesky decomposition of the $2 \times 2$ correlation matrix. Because the transition density is exact, the simulation causes no discretization error regardless of the step size $\Delta t$.

Yields are computed at each time step from the simulated factors using (5)–(7), with the calendar time $t$ passed to $A(t, \tau)$ so that the deterministic drift is correctly reflected in bond prices throughout the simulation.

**Calibration.** The model has two groups of parameters: structural parameters $(a_x, a_y, \sigma_x, \sigma_y, \rho)$ that govern the stochastic dynamics, and deterministic parameters $(\ell_0, \gamma, \kappa)$ that position the yield curve.

The structural parameters $a_x, a_y$, and $\rho$ are set by the user based on prior views about shock persistence and cross-maturity co-movement. The calendar-time trend $\gamma$ is set directly (default 0) as well.

Given these choices, the deterministic parameters $\ell_0$ and $\kappa$ are calibrated analytically from two target conditions on the initial yield curve:

1. The expected 2-year yield: $\mathbb{E}[y(0,2)] = \bar{y}_2$.
2. The expected 10y–2y spread: $\mathbb{E}[y(0,10)] - \mathbb{E}[y(0,2)] = \bar{s}_{10-2}$.

At $t$=0 with $x_0 = y_0 = 0$, the yield (9) gives

$$\mathbb{E}[y(0,\tau)] \;=\; \ell_0 + \tfrac{1}{2}(\gamma + \kappa)\,\tau - c(\tau), \tag{13}$$

where $c(\tau) = V(\tau)/(2\tau)$ is the convexity correction, computable in closed form from $a_x, a_y, \sigma_x, \sigma_y, \rho$. The spread condition determines the total slope $\gamma + \kappa = [\bar{s}_{10-2} + c(10) - c(2)]\,/\,4$, from which $\kappa = (\text{total slope}) - \gamma$. The level condition then gives $\ell_0 = \bar{y}_2 + c(2) - \gamma - \kappa$. This two-equation system has a unique closed-form solution for any choice of structural parameters and $\gamma$.

**Volatility calibration via negative-rate probability.** When $\sigma_x$ is not specified directly, it can be calibrated from a constraint on the probability of negative short rates. At horizon $T$, the short rate $r_T$ is Gaussian with

$$\mathbb{E}[r_T] = \ell_0 + \gamma T, \qquad \mathrm{Var}(r_T) = \frac{\sigma_x^2}{2a_x}(1 - e^{-2a_x T}) + \frac{\sigma_y^2}{2a_y}(1 - e^{-2a_y T}) + \frac{2\rho\,\sigma_x\sigma_y}{a_x + a_y}(1 - e^{-(a_x+a_y)T}). \tag{14}$$

The probability $\mathbb{P}(r_T < 0) = \Phi(-\mathbb{E}[r_T]/\sqrt{\mathrm{Var}(r_T)})$ is monotonically increasing in $\sigma_x$. Given a target probability (e.g. 5% at $T$=10 years), $\sigma_x$ is solved numerically, after which $\ell_0$ and $\kappa$ are recalibrated as above. This provides an intuitive alternative to specifying volatility directly: the user sets how likely negative rates are estimated to be, and the model infers the required dispersion.

## 3.5 Growth modeling

### 3.5.1 Introduction to the model in GDP forecasting

The Markov Switching Regime Forecaster (MSRF) class presented here introduces a hybrid two-regime Markov Switching Autoregressive (MS-AR) model with external regressors (interest rate changes). This approach allows the model to:

- Identify distinct economic regimes based on Finnish historical GDP growth and interest rate movements.
- Incorporate external shocks (for example monetary policy changes) via regime-specific sensitivities to interest rates.
- Generate probabilistic forecasts using Monte Carlo simulations, providing a distribution of possible GDP growth paths rather than point estimates.

### 3.5.2 Regime-Specific functions and model structure

The MSRF implements a two-regime MS-AR model with the following state-dependent equations:

$$g_t = b_{0,s_t} + \phi_{s_t} y_{t-1} + \gamma_{2,s_t} \Delta \mathrm{IR}_{2Y,t-2} + \gamma_{5,s_t} \Delta \mathrm{IR}_{5Y,t-2} + \gamma_{10,s_t} \Delta \mathrm{IR}_{10Y,t-2} + \epsilon_{s_t}, \tag{15}$$

$$\epsilon_{s_t} \sim N(0, \sigma^2_{s_t}), \tag{16}$$

where

- $g_t$ is the differences GDP growth at time $t$,
- $s_t \in \{0, 1\}$ is the unobserved state (Regime 0: Expansion-like; Regime 1: Recession-like),
- $b_{0,s_t}$ is the regime-specific intercept,
- $\phi_{s_t}$ captures auto-regressive persistence,
- $\gamma_{k,s_t}$ are the regime-specific sensitivities to interest rate changes for maturities $k \in \{2, 5, 10\}$,
- $\epsilon_{s_t}$ is the regime-specific error term.

The transition between regimes is governed by a transition matrix $P$:

$$P = \begin{bmatrix} p_{00} & 1 - p_{11} \\ 1 - p_{00} & p_{11} \end{bmatrix},$$

where $p_{00}$ and $p_{11}$ are the probabilities of staying in Regimes 0 and 1, respectively.

### 3.5.3 Hyperparameter optimization via conditional probability maximization

The MSRF class optimizes 14 hyperparameters using maximum likelihood estimation (MLE) under the conditional probability maximization rule. The optimization process is implemented as follows.

**Data preprocessing**
GDP levels are converted to log differences and scaled to percentage growth:

$$\Delta \log(\mathrm{GDP}_t) = 100 \times [\log(\mathrm{GDP}_t) - \log(\mathrm{GDP}_{t-1})].$$

Interest rate series (2Y, 5Y, 10Y) are differenced to derive basis point changes:

$$\Delta \mathrm{IR}_{k,t} = \mathrm{IR}_{k,t} - \mathrm{IR}_{k,t-1}, \quad k \in \{2, 5, 10\}.$$

All series are aligned to a quarterly frequency to avoid mixed-frequency noise.

**Log-likelihood function (objective function)**
The negative log-likelihood function is defined as:

$$-\mathcal{L} = -\sum_{t=1}^{T} \log \left( \sum_{s=0}^{1} \pi_{s,t} \cdot f(y_t | \mu_s, \sigma_s) \right),$$

where

- $\pi_{s,t}$ is the predicted probability of being in state $s$ at time $t$ (using the transition matrix $P$),
- $f(y_t|\mu_s, \sigma_s)$ is the regime-specific normal density function for the observed GDP growth $y_t$,
- $\mu_s = b_{0,s} + \phi_s y_{t-1} + \sum_k \gamma_{k,s} \Delta \mathrm{IR}_{k,t-2}$ is the regime-specific mean.

**Optimization bounds and initial guess**
The optimization is performed using the COBYLA algorithm (a derivative-free method suitable for bounded problems). The COBYLA algorithm is suitable for this use case since our objective is smooth non-convex problem, which includes constraints, and COBYLA is also robust to poor initial guesses. The bounds are set to ensure economic plausibility. The parameter bounds for optimization are set to be equal for both regime's to avoid arbitrary asymmetry, meaning there was not strong evidence based on historical data that regimes have significantly different bounds. Thus, when parameters of the model are optimized for regimes they differ from each other, which resulted to one volatile and slowly growing regime, and for one steadily growing regime:

**Table 2:** Parameter bounds for optimization

| **Parameter** | **Regime 0** | **Regime 1** |
|---|---|---|
| Intercept ($b_0$) | None | None |
| Persistence ($\phi$) | $[-0.75, 0.8]$ | $[-0.75, 0.8]$ |
| Sensitivity to interest rates ($\gamma_k$) | None | None |
| Volatility ($\sigma$) | $[0.1, 5.0]$ | $[0.1, 5.0]$ |
| Transition probabilities ($p$) | $[0.1, 0.9]$ | $[0.1, 0.9]$ |

#### 3.5.4 Stochastic Monte Carlo simulations using G2++ model forecasted interest rate data

The MSRF class provides a method, `forecast_stochastic`, to generate Monte Carlo simulations of future GDP growth paths.

**Regime uncertainty**

Future states are simulated using the transition matrix $P$:

$$\text{Regime at } t+1 \sim P(\text{Regime}_i|\text{Regime}_t).$$

**Shock propagation using forecasted interest rates**

The model uses forecasted interest rate data (2Y, 5Y, 10Y) generated by the G2++ term structure model. The forecasted changes in interest rates are fed into the MSRF as external regressor:

$$\Delta\text{IR}_{2Y,t}, \Delta\text{IR}_{5Y,t}, \Delta\text{IR}_{10Y,t} \qquad \text{(forecasted data)}$$
$$f(y_t|\mu_s, \sigma_s) \qquad \text{(regime-specific mean)}$$

The forecasted interest rate paths are prepared as follows:

- The last known level of each interest rate series is prepended to calculate the first future difference.
- The differences are stacked into a matrix for use as external regressors in the stochastic simulations.

**Monte Carlo simulation logic**

For each simulation ($n = 50$):

1. The current state probabilities are filtered using the historical data and transition matrix $P$.
2. The next regime is sampled based on the transition probabilities.
3. A new GDP growth value is generated based on the current regime's autoregressive equation and the forecasted interest rate changes.
4. This process is repeated for all future time steps.

The method returns:

- An array of 50 simulated GDP growth paths (as a default),
- The median forecast (50th percentile),
- The 25th and 75th percentiles for uncertainty quantification.

### 3.5.5 Other GDP modeling methods

**Vector autoregression (VAR)**

Vector Autoregression (VAR) models capture linear inter-dependencies between macroeconomic variables such as GDP growth, inflation, and interest rates using a system of equations (Lütkepohl 2005). It is widely used for policy simulations due to its ability to quantify impulse responses and forecast error variance decompositions. However, VAR assumes constant parameters and does not account for regime shifts, limiting its ability to model nonlinear dynamics in GDP growth.

**Bayesian vector autoregression (BVAR)**

Bayesian Vector Autoregression (BVAR) extends VAR by incorporating prior distributions for parameters, enabling probabilistic forecasting and improved uncertainty quantification (Giannone and Reichlin 2006). BVAR is particularly useful for risk assessment and scenario analysis. However, it retains the linearity assumption of VAR and does not explicitly model regime shifts.

**Machine learning models**

Machine learning models (e.g., Random Forest, XGBoost, neural networks) are effective at capturing nonlinear relationships and feature importance in GDP growth data (Breiman 2001). However, they operate as "black boxes," offering limited interpretability, theoretical insights, and transparency.

**Structural macroeconomic models**

Structural macroeconomic models (e.g., DSGE) offer theoretical insights into economic mechanisms and are designed for policy simulations (Smets and Wouters 2003). However, they are computationally intensive, require extensive calibration, and lack flexibility in adapting to real-world data and regime shifts.

### 3.5.6 Final summary

The Markov Switching Regime Forecaster (MSRF) represents a transparent framework for modeling GDP growth under interest rate uncertainty. By capturing regime shifts and state-dependent dynamics, it provides insight into how economic conditions can evolve and respond to external shocks, such as changes in interest rates. The model's ability to generate probabilistic forecasts through Monte Carlo simulations aligns some of the initial client preferences - Model is able to provide more than one scenario, with normally distributed stochastic error term and it is interconnected with forecasted interest rates. Thus, model is less transparent as it could be (for example when comparing to time series models).

However, the MSRF is sensitive to changes in datasets and outlier interest rate scenarios. The optimization process, which relies on the COBYLA solver to maximize the log-likelihood function, can produce varying results depending on the input data's quality and the presence of extreme values in interest rate forecasts. For instance, sudden spikes or prolonged deviations in interest rate paths (for example untypical monetary policy shocks) may lead to regime misclassifications or exaggerated volatility estimates in the forecasts. Similarly, structural breaks or shifts in the underlying economic relationships (for example post-pandemic regimes) can impact the model's ability to accurately identify and transition between states.

## 3.6 Optimization model (debt strategy under uncertainty)

This section describes the optimization layer for selecting an issuance strategy under macro-financial uncertainty. The formulation follows the scenario-based sovereign debt risk-management perspective in Zenios et al. (2021), but scenarios are generated by stochastic models (G2++ model for interest rates and a Markov regime model for GDP) rather than an explicit scenario tree (Brigo and Mercurio 2006; Hamilton 1989; Kim and Nelson 1999). The objective is to minimise expected interest expenditures while controlling funding and sustainability risks using coherent tail-risk constraints. This is aligned with the practical principles of public debt management, where the maturity structure is chosen to balance expected cost with robustness under adverse scenarios (International Monetary Fund and World Bank 2014; Missale 1999; Belton et al. 2018).

### 3.6.1 Scenario set and inputs

Let $n \in \{1, \dots, N\}$ index scenarios and $t \in \{1, \dots, T\}$ index time periods. Each scenario provides a path for (i) yield-curve variables needed to price new issuance and refinancing, (ii) GDP $Y_{n,t}$, and (iii) the primary balance $PB_{n,t}$. Scenario probabilities are denoted by $p_n$ with $\sum_{n=1}^{N} p_n = 1$ (baseline case: $p_n = 1/N$). Although multistage stochastic programming is often formulated using scenario trees, Monte Carlo style scenario sets can be used provided numerical stability of tail estimates is checked as $N$ increases (Høyland and Wallace 2001).

### 3.6.2 Decision variables and strategy parametrization

The state is conditioned on an inherited (legacy) debt stock that cannot be changed. The decision is the new issuance allocation across $J$ maturity buckets (or instrument types). A fully scenario- and time-adaptive issuance decision $X_{n,t,j}$ is conceptually possible but can be overly complex and difficult to validate within a policy context. Therefore, the baseline implementation uses transparent strategy parameterizations:

**Time-invariant fixed mix.** New issuance is proportional to financing needs with constant weights

$$X_{n,t,j} = w_j \cdot GFN_{n,t}, \qquad \sum_{j=1}^{J} w_j = 1, \qquad w_j \geq 0. \tag{17}$$

**Adaptive fixed mix (block-wise).** Weights are allowed to change across a small number of blocks $b(t) \in \{1, \dots, B\}$

$$X_{n,t,j} = w_{b(t),j} \cdot GFN_{n,t}, \qquad \sum_{j=1}^{J} w_{b,j} = 1, \qquad w_{b,j} \geq 0. \tag{18}$$

These parameterizations are common because they lower the risk of overfitting, improve interpretability, and make it easier to communicate strategies as maturity shares (Belton et al. 2018;

Missale 1999).

### 3.6.3 Debt accounting, financing need, and sustainability

We define gross financing need (GFN) as the cash requirement that must be funded by issuance. Following the sovereign debt accounting logic used in scenario-based frameworks (Zenios et al. 2021; Bolder 2003), and write

$$GFN_{n,t} = I_t^{legacy} + A_t^{legacy} - PB_{n,t} + I_{n,t}^{new} + A_{n,t}^{new}, \tag{19}$$

where $I_t^{legacy}$ and $A_t^{legacy}$ are exogenous interest and amortization flows from inherited debt, and $I_{n,t}^{new}$ and $A_{n,t}^{new}$ are endogenous flows implied by issuance decisions and scenario-dependent yields.

Debt stock evolution is represented by a simplified stock-flow identity

$$Debt_{n,t} = Debt_{n,t-1} + GFN_{n,t} - A_t^{legacy} - A_{n,t}^{new}. \tag{20}$$

A sustainability cap can be imposed as a hard constraint

$$Debt_{n,t} \leq d^{\max} Y_{n,t}, \tag{21}$$

where $d^{\max}$ is a policy parameter that sets the cap on how much debt can be carried with respect to the GDP. This is consistent with standard public debt dynamics discussions and sustainability analysis (Escolano 2010; International Monetary Fund and World Bank 2014).

## 3.7 Objective function

The baseline objective minimizes expected net interest payments (NIP) over the horizon

$$\min \sum_{n=1}^{N} p_n \sum_{t=1}^{T} NIP_{n,t}, \qquad NIP_{n,t} = I_t^{legacy} + I_{n,t}^{new}. \tag{22}$$

This matches the principle in (Zenios et al. (2021)): expected cost is minimized subject to explicit risk controls and sustainability constraints.

## 3.8 Risk constraints (flow and stock tail risk)

To control downside outcomes, the model constrains tail risk using Conditional Value-at-Risk (CVaR), which is a coherent risk measure with a convex optimization representation (Rockafellar, Uryasev, et al. 2000). The framework implements two families of risk constraints:

**Flow risk (CFaR-type) on funding pressure.** Define the funding-pressure ratio

$$gfn_{n,t} = \frac{GFN_{n,t}}{Y_{n,t}}. \tag{23}$$

For each $t$, CVaR at confidence level $\alpha$ is constrained by a policy cap $\omega_t$

$$CVaR_\alpha(gfn_{\cdot,t}) \leq \omega_t. \tag{24}$$

Using the standard Rockafellar–Uryasev linearization, introduce an auxiliary variable $\eta_t$ and slack variables $u_{n,t} \geq 0$:

$$u_{n,t} \geq gfn_{n,t} - \eta_t, \tag{25}$$

$$u_{n,t} \geq 0, \tag{26}$$

$$\eta_t + \frac{1}{1-\alpha} \sum_{n=1}^{N} p_n u_{n,t} \leq \omega_t. \tag{27}$$

**Stock risk (CDeaR-type) on debt growth.** Debt growth is governed by

$$\Delta Debt_{n,t} = Debt_{n,t} - Debt_{n,t-1}. \tag{28}$$

A CVaR-type cap $\bar{\omega}_t$ puts an upper bound on the expected leve of extreme debt

$$CVaR_\alpha(\Delta Debt_{\cdot,t}) \leq \bar{\omega}_t, \tag{29}$$

with the analogous linearization using $\xi_t$ and nonnegative slacks $v_{n,t}$

$$v_{n,t} \geq \Delta Debt_{n,t} - \xi_t, \tag{30}$$

$$v_{n,t} \geq 0, \tag{31}$$

$$\xi_t + \frac{1}{1-\alpha} \sum_{n=1}^{N} p_n v_{n,t} \leq \bar{\omega}_t. \tag{32}$$

These constraints enforce the key principle that policy must be robust to adverse scenarios by limiting tail outcomes in both funding pressure and debt accumulation (Zenios et al. 2021; Rockafellar, Uryasev, et al. 2000).

### 3.8.1 Computational properties and implementation

Under fixed-mix strategy parametrization, the optimization remains tractable because the decision space is low dimensional (weights $w$), and the CVaR constraints admit linear reformulations (Rockafellar, Uryasev, et al. 2000). The model is implemented in Pyomo and solved with a linear programming solver (HiGHS). This make it possible to carry out repeated experiments across portfolios, scenario sets, and risk-cap configurations. The main practical challenges are related not to solver feasibility but to modeling choices: scenario representativeness, calibration of macro regimes, and stability of tail-risk estimates as the scenario count changes (Bolder 2003; Høyland and Wallace 2001).

Finally, the optimization result is reported through a decision-oriented interface: strategy weights by maturity bucket, issuance profiles, and horizon-dependent cost and sustainability metrics.

This matches the emphasis in public debt management practice on interpretable strategy levers and robust performance under stress (International Monetary Fund and World Bank 2014; Belton et al. 2018).

# 4 Results

This section presents some results of the based framework applied to one of the simulated states' debt situation. The results are organized into two sections: (i) A presentation of the baseline GDP and interest forecast and the parameters for that were used in their calibration, (ii) presentation of the different metrics relating to each portfolio we compare.

## 4.1 Baseline scenario forecast

Under the baseline scenario with parameters given in Tables in 3 and 4, the framework provides forecasts for GDP and interest rates shown in Figures 1a and 1b. The GDP is estimated to grow from the initial 300 bn euros to around 1 400 billion at the end of our horizon (2056). The interest rates are assumed to have a slight term premium and some volatility with the median path of each shown in Fig. 1b We chose three different portfolios for the analysis to illustrate and compare the risk cost tradeoff. Each portfolio represents a general strategy of either holding mainly short or long debt or equal amount of all three maturities. The portfolio maturity weights are presented in Table 5.

**Table 3:** Baseline macro-fiscal and debt portfolio parameters.

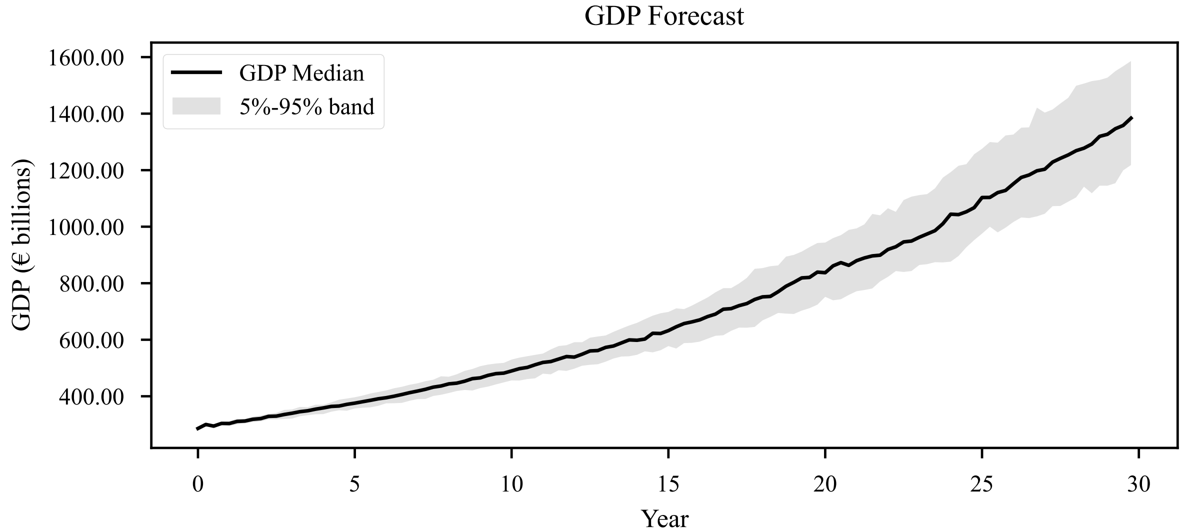
| Parameter | Value | Description |
|---|---|---|
| $H$ | 30 years | Simulation horizon |
| $\Delta t$ | 0.25 years | Period length (quarterly) |
| $N$ | 50 | Number of stochastic scenarios |
| $GDP_0$ | 300 bn EUR | Initial nominal GDP |
| $PB_0$ | -1.0 bn EUR | Initial per-period primary balance |
| $Debt_0$ | 190 bn EUR | Initial central government debt |
| $Cash_0$ | 6 bn EUR | Initial cash buffer |
| $\mu_g$ | 2.0% | Annual real GDP growth baseline |
| $\alpha$ | 95% | CVaR confidence level for risk metrics |
| Maturities | {2, 5, 10} years | Evaluated issuance instruments (Buckets) |

**Table 4:** Calibrated G2++ (decoupled) stochastic interest rate parameters.

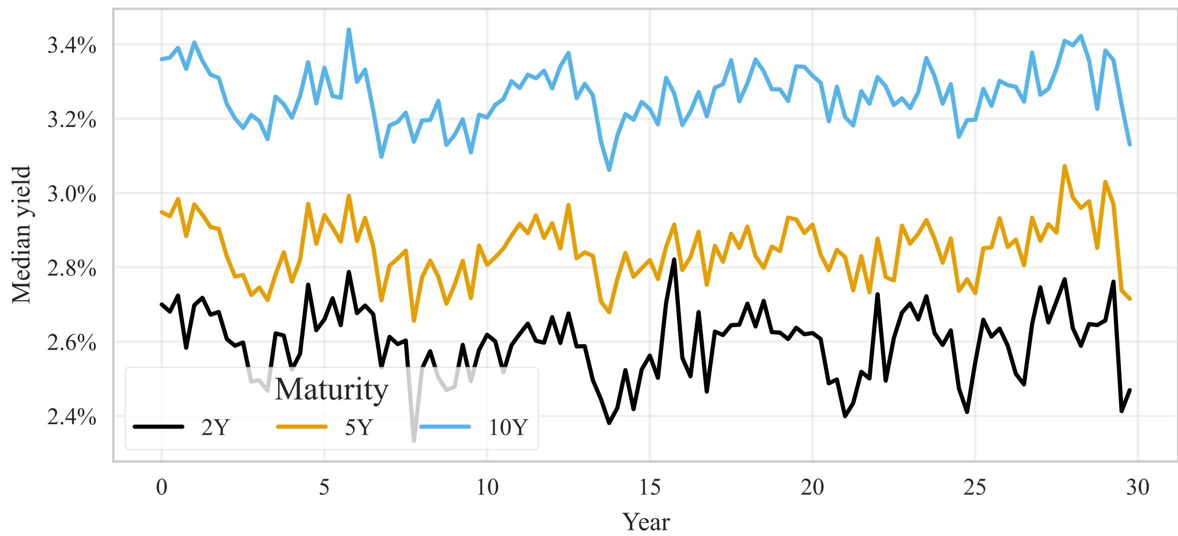
| Parameter | Symbol | Value |
|---|---|---|
| Mean-reversion speeds | $a_x,\ a_y$ | 0.50, 0.03 |
| Factor volatilities | $\sigma_x,\ \sigma_y$ | 0.006, 0.003 |
| Correlation | $\rho$ | 0.10 |
| Baseline level | $\ell_0$ | 2.53% (calibrated Level0) |
| Calendar-time trend | $\gamma$ | 0.00% (Stationary long-run) |
| Slope adjustment | $\kappa$ | 0.17% (calibrated for 10y-2y spread) |

**Table 5:** Issuance allocation weights across maturity buckets for baseline and alternative strategies.

| Strategy | 2-Year | 5-Year | 10-Year |
|---|---|---|---|
| Baseline (Even) | 33.3% | 33.3% | 33.3% |
| Short | 80.0% | 15.0% | 5.0% |
| Long | 5.0% | 15.0% | 80.0% |




**(a)** GDP forecast of the baseline scenario




**(b)** Evolution of bond yields in the baseline scenario. The median path is displayed.

**Figure 1:** Baseline scenarios for GDP forecast and interest rate dynamics.

## 4.2 Debt cost for risk under baseline scenario for different strategies

Under the above presented scenario the median GFN (Gross Financing Need) to GDP (Gross Domestic Product) ratio evolves as in Fig. 2. The largest financing need arises from using the short portfolio and the lowest from the long portfolio. The median issuance profile in Fig. 6 is the lowest in the long portfolio and highest in the even one with the short in between the two extremes.

The debt-to-GDP ratio in Fig. 3 shows a declining trend. The short portfolio seems to outperform the other two slightly towards the end of the horizon.

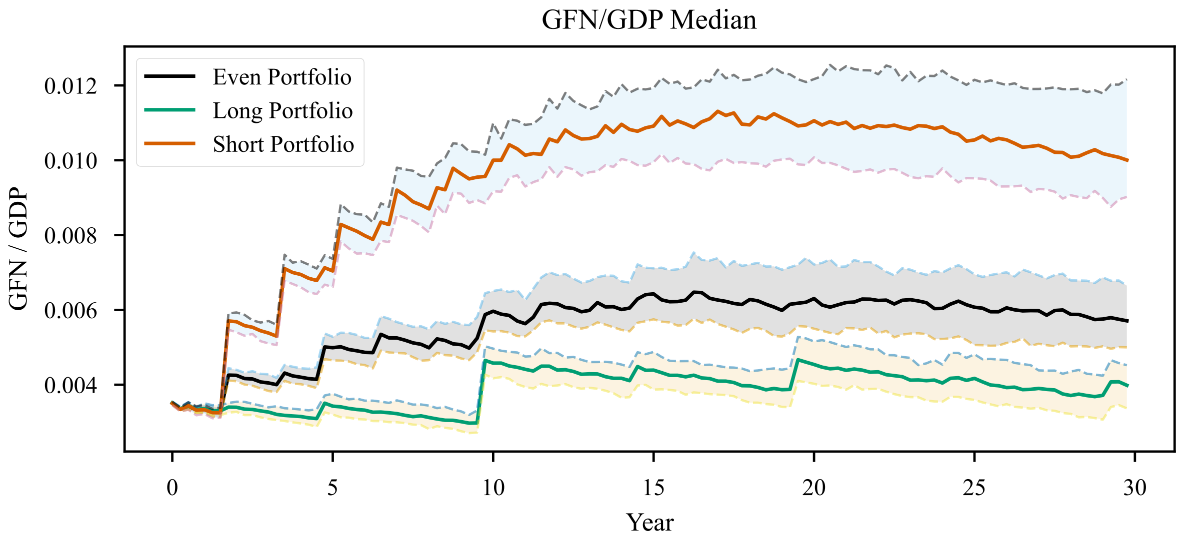



**Figure 2:** Median GFN/GDP.

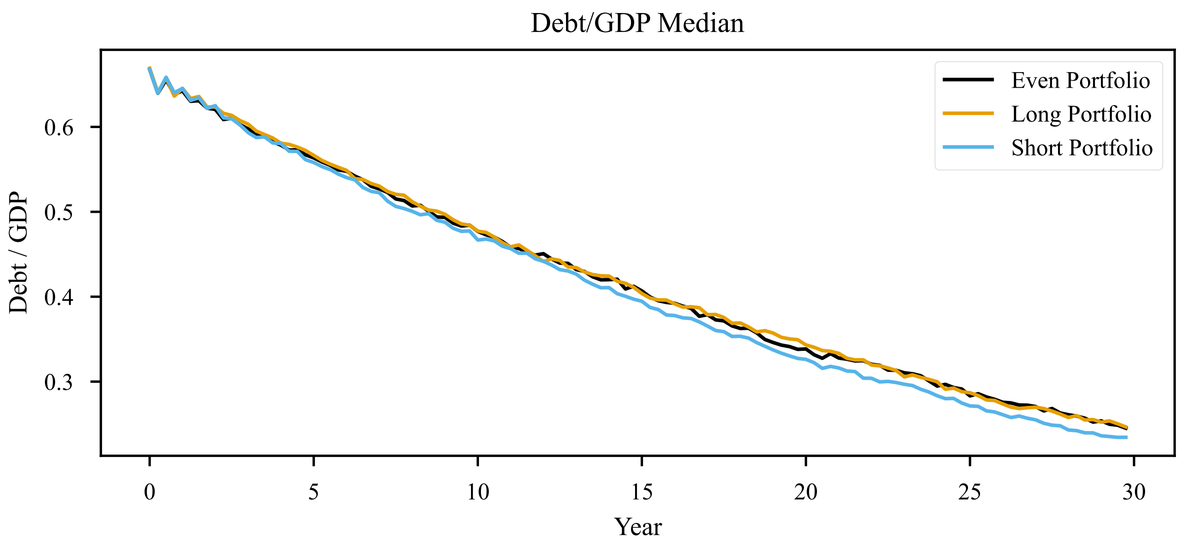
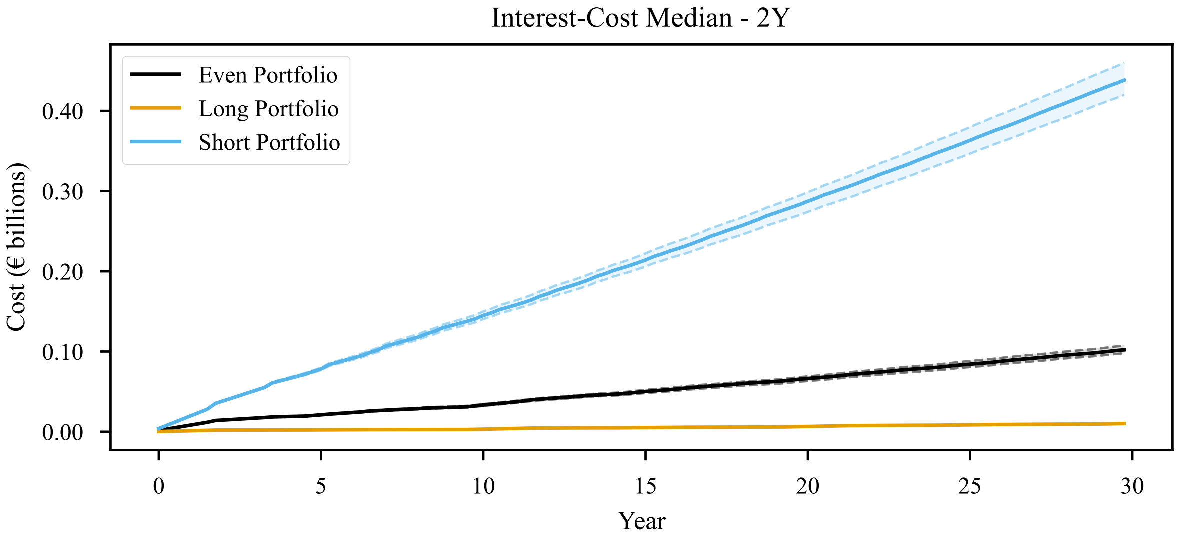
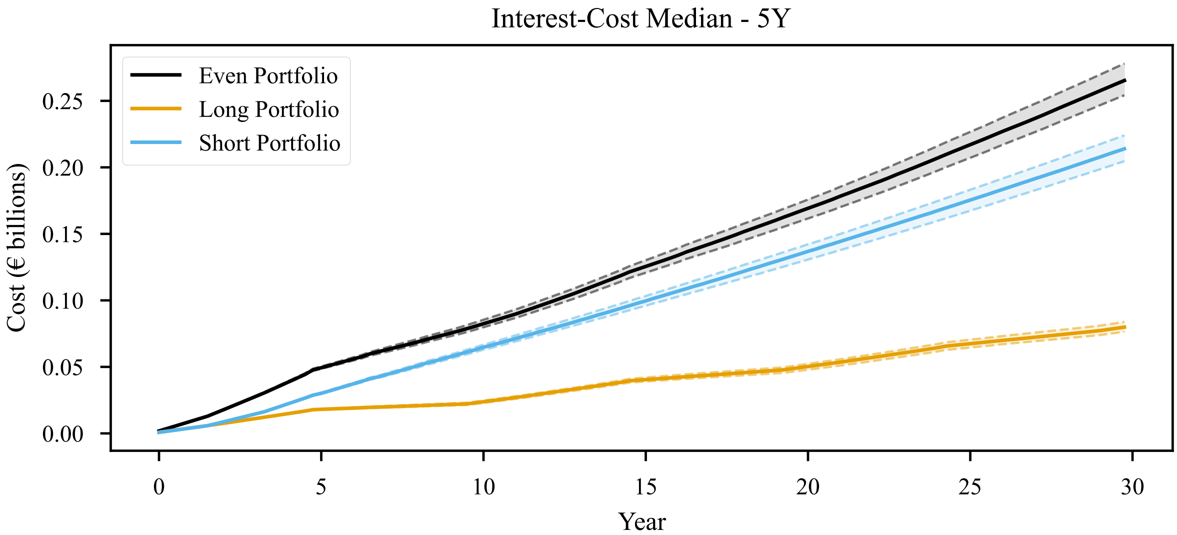


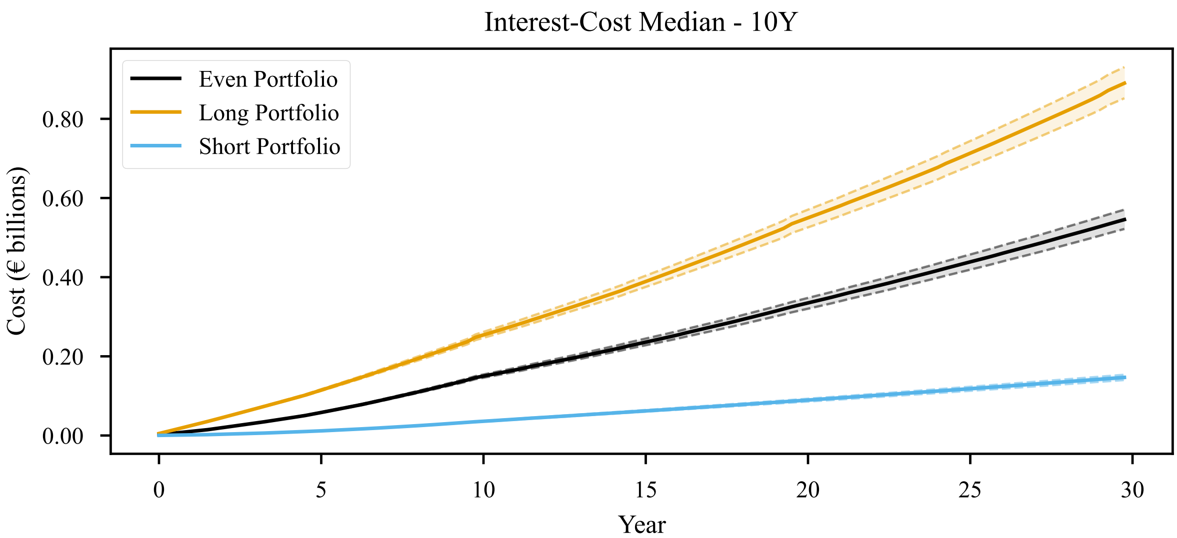

**Figure 3:** Median debt-to-GDP ratio.




**(a)** Median interest cost (2-year horizon)




**(b)** Median interest cost (5-year horizon




**(c)** Median interest cost (10-year horizon)

**Figure 4:** Median interest cost comparison across all horizons.

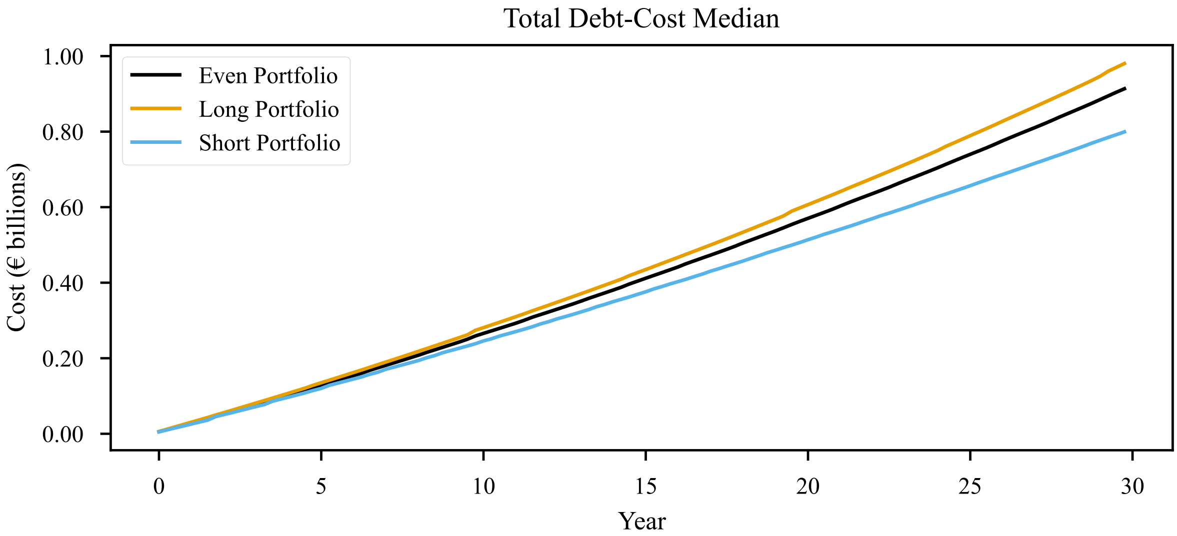



**Figure 5:** Total debt cost median in each scenario.

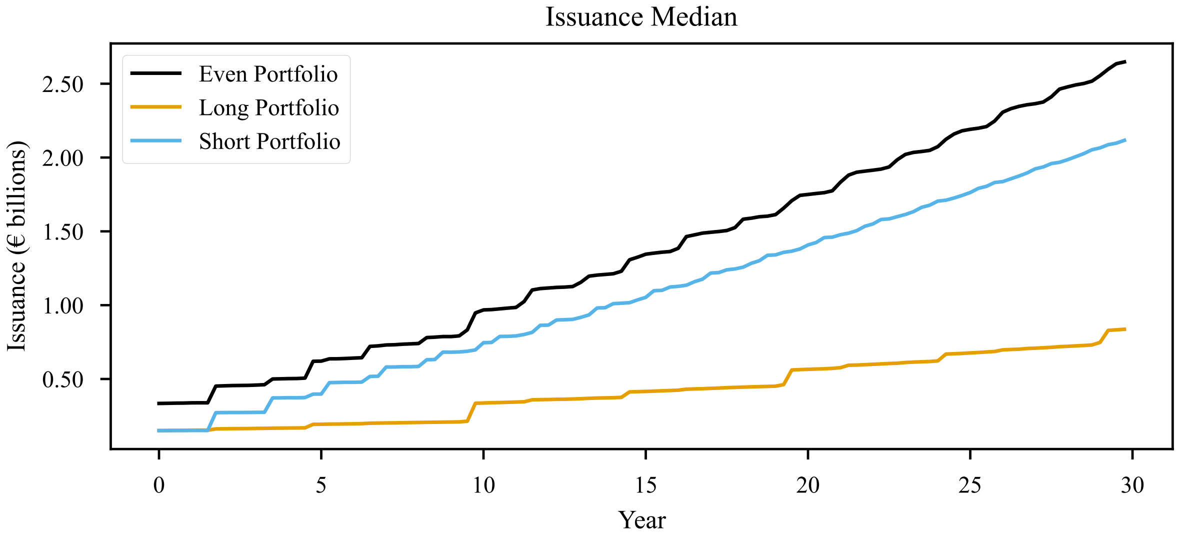



**Figure 6:** Median issuance profile.

# 5 Discussion

## 5.1 Interpretation of the results

Median interest cost comparisons at multiple horizons (Figures 4a–4c) show that strategy rankings can change with the horizon, which is consistent with the fact that refinancing exposure accumulates over time. Median Debt/GDP outcomes (Figure 3) and median GFN/GDP outcomes (Figure 2) translate financing costs and refinancing profiles into sustainability and funding-risk implications. Finally, the strategy levers are visible in the issuance profile (Figure 6) and in the maturity-bucket allocation weights (Figure 5).

A central qualitative conclusion is that the relevant decision is not whether one strategy minimizes cost in the median. The relevant decision is which strategy provides an acceptable balance between (i) typical funding cost, (ii) refinancing pressure and financing flexibility, and (iii) robustness under adverse rate and growth regimes. This cost–risk trade-off is emphasized in the sovereign debt strategy literature (Missale 1999; Abbas, Pienkowski, and Rogoff 2019; Belton et al. 2018).

## 5.2 Key risk drivers and cost–risk trade-offs

**Refinancing profile and maturity concentration.** A primary driver of both funding risk and interest cost dispersion is the refinancing profile implied by issuance. Even when inherited debt is fixed, new issuance decisions determine how rapidly the portfolio rolls over. A strategy that concentrates issuance in short maturities increases the frequency of refinancing and thereby increases sensitivity to rate shocks. This channel is visible in the issuance profile and the issuance weights. It is also reflected in horizon-dependent cost comparisons: differences that are modest at two years can become material by ten years once refinancing exposure accumulates (Figures 4a–4c).

**Yield-curve level and slope dynamics.** A second driver is the evolution of the yield curve. Changes in curve level and slope alter both marginal issuance costs and the effective refinancing rate. This is visible in the refinancing-rate comparison (Figure 6) and indirectly in the cost horizons (Figures 4a–4c). A multi-factor rate model is therefore essential because a single-factor specification often fails to represent the joint short- and long-end dynamics that matter for portfolio cash flows. This is one reason affine multi-factor term-structure models are widely used as scenario engines (Brigo and Mercurio 2006; Duffie and Kan 1996; Dai and Singleton 2000).

**Macroeconomic persistence and compounding.** Debt sustainability is strongly influenced by GDP dynamics, not only by rates. A persistent low-growth regime increases Debt/GDP mechanically and amplifies the fiscal effect of higher refinancing rates through an adverse interest–growth differential. This mechanism is consistent with standard debt-dynamics decompositions and sustainability analysis (Escolano 2010). In the results, the GDP median path (Figure 1a) drives the denominator of Debt/GDP, which then influences Debt/GDP (Figure 3) and GFN/GDP (Figure 2). The Markov regime assumption matters because it introduces persistence: adverse regimes can last long enough for compounding effects to dominate (Hamilton 1989; Kim and Nelson 1999).

**Constraints and implementation feasibility.** A fourth driver is feasibility and implementation constraints. Restrictions that smooth issuance over time, limit maturity concentration, or enforce maturity-structure targets can reduce refinancing pressure and limit undesirable issuance profiles. Such constraints are common in practice and appear in model-based Treasury analyses (Belton et al. 2018; Missale 1999). However, constraints can increase expected costs by restricting the model's ability to exploit short-term cheap funding. The issuance profile and weights (Figure 6) therefore reflect not only chosen strategies but also the binding structure of the decision problem.

## 5.3 Horizon dependence

The horizon comparisons (Figures 4a–4c) highlight that strategy rankings can be horizon dependent even when all strategies are feasible. Short-horizon differences are dominated by near-term refinancing and the part of the curve that prices near-term issuance. Long-horizon outcomes incorporate multiple refinancing cycles, regime persistence in GDP, and the cumulative effect of the refinancing rate path. As a result, a strategy that is cost-effective in the near term can become less attractive over horizons that are more relevant for sustainability and long-run robustness.

This observation supports a practical recommendation: strategy evaluation should be performed at multiple horizons and should report both typical and adverse outcomes. This is consistent with public debt management guidance and model-based Treasury frameworks that emphasize robustness across plausible futures (International Monetary Fund and World Bank 2014; Belton et al. 2018).

## 5.4 Key model assumptions

**Interest-rate dynamics (G2++ model).** The G2++ model is a two-factor Gaussian model with an affine structure and a deterministic shift that matches the initial yield curve (Brigo and Mercurio 2006). The most influential parameters are mean-reversion speeds, factor volatilities, and the correlation between factors. These govern persistence of rate shocks, dispersion of yield scenarios, and cross-maturity covariance. Because strategies load differently on maturities, these parameters can change both the level and ordering of horizon cost comparisons (Figures 4b–4c). Sensitivity analysis with respect to these parameters and explicit stress overlays for abrupt curve shifts are therefore necessary for credible conclusions.

**GDP dynamics (Markov regimes).** The GDP engine assumes regime switching, which implies persistent growth states governed by a transition matrix. Transition probabilities and regime-specific growth moments are influential because they determine how long adverse states last and how rapidly the economy transitions back to a favourable regime (Hamilton 1989; Kim and Nelson 1999). Higher persistence in low-growth regimes amplifies debt dynamics and can materially worsen Debt/GDP and GFN/GDP outcomes (Figures 3–2). These parameters should therefore be treated as primary uncertainty drivers.

**Rate–growth dependence.** Sustainability outcomes can be dominated by whether adverse growth regimes coincide with elevated refinancing rates. If the model assumes low growth coincides with higher term premia or higher refinancing rates, tail risk increases. Empirical evidence on the direction of interest-rate and GDP-growth relationships can be context dependent (Lee and Werner 2018). The dependence structure should therefore be documented transparently and tested under alternative specifications.

**Objective function and risk definition (expected cost vs tail risk).** The chosen objective materially affects conclusions. A pure expected-cost objective tends to favor short maturities, whereas tail-risk constraints or penalties shift solutions toward smoother refinancing profiles and more robust outcomes. A coherent tail-risk measure is Conditional Value-at-Risk (CVaR), which has a convex optimization representation (Rockafellar, Uryasev, et al. 2000):

$$CVaR_{\alpha}(L) = \eta + \frac{1}{1-\alpha}\mathbb{E}\big[(L-\eta)^{+}\big].$$

In a sovereign setting, $L$ can be defined as interest expenditure, GFN/GDP, or a sustainability loss metric, while $\alpha$ represents risk tolerance. This is consistent with scenario-based sovereign debt frameworks that manage tail risk explicitly (Zenios et al. 2021).

### 5.5 Model validity and limitations

The framework is credible as decision support because it generates coherent interest-rate scenarios using a standard term-structure model (Brigo and Mercurio 2006), captures macro persistence via regime switching (Hamilton 1989; Kim and Nelson 1999), and maps these risk factors into debt cash flows and strategy constraints in a scenario-based evaluation (Bolder 2003; Belton et al. 2018). However, limitations must be stated explicitly.

- **Gaussian rate dynamics and extreme events.** The G2++ model is Gaussian and may underrepresented asymmetric stress behavior, liquidity events, or sudden risk-premium jumps. Stress overlays partially address this limitation, but they do not substitute for explicit jump or regime-switching rate models (Brigo and Mercurio 2006).
- **Reduced-form macro modeling.** Markov regime GDP modeling is stylized. Transition probabilities can be unstable under structural change and regime definitions can be sample dependent (Hamilton 1989; Kim and Nelson 1999).
- **Dependence assumptions and model risk.** Severe sustainability outcomes can be driven by the assumed co-movement of growth and refinancing rates. Misclassifications can change strategy rankings (Escolano 2010; Lee and Werner 2018).
- **Scenario approximation and numerical stability.** Tail-risk estimates can be sensitive to scenario count and discretizations. Scenario-tree methods are an alternative multistage approximation but introduce their own construction challenges (Høyland and Wallace 2001).

- **Institutional and execution feasibility.** Optimization can recommend strategies that are mathematically optimal under the model but not fully implementable due to issuance conventions, market capacity, or policy constraints. Including explicit constraints improves realism, but results should be interpreted as structured decision support rather than a mechanical prescription (Missale 1999; Greenwood et al. 2014; Belton et al. 2018).

## 5.6 Implications for decision-making

The results support a practical conclusion: expected-cost minimization alone is not sufficient for sovereign debt strategy. Decision-making should incorporate funding and sustainability risk metrics (Figures 2 and 3) and evaluate strategies over multiple horizons (Figures 4a– 4c). A transparent cost–risk comparison, supported by CVaR-type objectives or constraints, offers a principled way to align issuance strategy selection with explicit risk tolerance (Rockafellar, Uryasev, et al. 2000; Zenios et al. 2021). This approach is consistent with public debt management guidance and model-based Treasury frameworks that emphasize robustness and stress testing (International Monetary Fund and World Bank 2014; Belton et al. 2018; Bolder 2003).

Overall, the framework is most valuable as a tool for comparing feasible strategies under uncertainty, quantifying downside risk transparently, and supporting policy discussion based on explicit assumptions and documented sensitivity analysis.

# 6 Conclusions

The objective of this project was to develop a simulation-based, decision-relevant framework that supports the State Treasury in ensuring the government's solvency at the lowest feasible cost, while controlling downside risk even under adverse macro-financial scenarios. In line with the risk-management view of sovereign financing, the framework produces distributions (not point forecasts) for key outcomes and enables explicit cost–risk trade-offs through optimization under constraints (Zenios et al. (2021)).

Methodologically, we implemented an end-to-end pipeline that combines (i) stochastic scenario generation for the term structure of interest rates (G2++ model), (ii) a Markov regime model for GDP dynamics, and (iii) a debt cash-flow and stock accounting engine linked to an optimization model. The optimization layer is implemented in Pyomo and can be solved efficiently using the HiGHS solver, enabling iterative experiments and strategy comparison across scenario sets. A key practical outcome is that the model produces interpretable outputs that match the needs of debt management: issuance profiles and issuance mix, debt-to-GDP paths, gross financing needs (GFN) relative to GDP, and interest-cost distributions by instrument and in total. These diagnostics are generated as risk bands (5%–50%–95% quantiles) to support tail-risk interpretation and communication.

A preliminary test run demonstrates that the full optimization and reporting pipeline is operational. The model solves successfully under HiGHS and reports an "expected net-interest" objective value of approximately $3.534 \times 10^{11}$ . Under the same run, the fixed-mix strategy weights (six maturity buckets) place the largest weight on a single bucket (0.55), with remaining weights distributed across the other maturities $(0.05, 0.15, 0.10, 0.10, 0.05)$ . Distributional sustainability and funding risk diagnostics from the same run include an expected debt ratio at the horizon end of 1.2334, a median debt ratio peak of 1.2264, and a median GFN/GDP peak of 0.1168 . In addition, the reporting layer produces period-by-period risk bands for Debt/GDP and GFN/GDP and visualizes instrument-level and total interest-cost distributions. These outputs are directly aligned with the decision context: they quantify both average performance and tail behavior.

The results reinforce a central conclusion from the sovereign debt literature: cost minimization and risk control are in tension, and the relevant ranking of strategies depends on the chosen risk tolerance and constraints. In the current implementation, tightening flow and stock risk caps (for example, CFaR-type constraints) can reduce refinancing and sustainability tail risk, but it may increase expected cost and, if pushed too far, lead to infeasibility. This is a practical policy insight: meaningful risk control requires explicit choices about acceptable trade-offs and feasible issuance dynamics.

The main limitations of the current results are related to calibration and structural assumptions. First, the numerical values reported above are conditional on the configuration used in the test run and should be interpreted as a demonstration of framework functionality rather than final calibrated policy conclusions . Second, the Gaussian rate dynamics of the G2++ model are tractable and coherent for yield-curve scenario generation, but they may not fully capture extreme non-Gaussian stress behavior without explicit stress overlays. Third, the Markov GDP model captures regime persistence but remains a stylized representation of macro dynamics, and tail outcomes can be sensitive to regime transition probabilities and assumed dependence structures. Finally, the optimization can be extended by enabling endogenous debt-rate feedback iterations, which are currently optional and must be switched on explicitly in the configuration.

The project delivers a functioning analytical framework that (i) generates coherent scenarios, (ii) maps them to debt-service and sustainability outcomes, and (iii) supports strategy comparison under explicit cost–risk objectives and constraints. The immediate next steps are calibrating and validating scenario engines against historical and client-relevant stylized facts, systematic sensitivity analysis with increasing scenario counts to stabilize tail estimates, structured experiments across issuance rules and risk-cap settings and implementing the endogenous feedback option where appropriate. These steps will convert the current operational prototype into a viable decision-support tool for the State Treasury.

## References

Abbas, S. Ali, Alex Pienkowski, and Kenneth Rogoff (Oct. 2019). *Sovereign Debt: A Guide for Economists and Practitioners*. Oxford University Press. ISBN: 9780198850823. DOI: 10.1093/oso/9780198850823.001.0001. URL: https://doi.org/10.1093/oso/9780198850823.001.0001.

Belton, Terry, David Greenlaw, Brian Sack, Huachen Li, Kristopher Dawsey, and Srini Ramaswamy (Oct. 2018). *Optimizing the Maturity Structure of U.S. Treasury Debt: A Model-Based Framework*. Hutchins Center Working Paper 46. Hutchins Center on Fiscal & Monetary Policy at Brookings. URL: https://www.brookings.edu/research/optimizing-the-maturity-structure-of-u-s-treasury-debt.

Bolder, David (Feb. 2003). "A Stochastic Simulation Framework for the Government of Canada's Debt Strategy". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.1082792.

Breiman, Leo (2001). "Random Forests". In: *Machine Learning* 45.2, pp. 5–32.

Brigo, Damiano. and Fabio. Mercurio (2006). *Interest Rate Models - Theory and Practice : With Smile, Inflation and Credit*. eng. 2nd ed. Springer finance. Berlin: Springer. ISBN: 3540221492.

Dai, Qiang and Kenneth J. Singleton (Oct. 2000). "Specification Analysis of Affine Term Structure Models". In: *The Journal of Finance* 55.5, pp. 1943–1978. DOI: 10.1111/0022-1082.00278.

Duffie, Darrell and Rui Kan (Oct. 1996). "A Yield-Factor Model of Interest Rates". In: *Mathematical Finance* 6.4, pp. 379–406. DOI: 10.1111/j.1467-9965.1996.tb00123.x.

Escolano, Julio (Jan. 2010). *A Practical Guide to Public Debt Dynamics, Fiscal Sustainability, and Cyclical Adjustment of Budgetary Aggregates*. Technical Notes and Manuals 10/02. International Monetary Fund, Fiscal Affairs Department. URL: https://www.imf.org/external/pubs/ft/tnm/2010/tnm1002.pdf (visited on 04/13/2026).

Giannone, Domenico and Lucrezia Reichlin (2006). "Forecasting in Real Time". In: *Quarterly Journal of Economics* 121.2, pp. 521–538.

Greenwood, Robin, Samuel G. Hanson, Joshua S. Rudolph, and Lawrence H. Summers (Sept. 2014). *Government Debt Management at the Zero Lower Bound*. Hutchins Center Working Paper 5. The Brookings Institution. URL: https://www.brookings.edu/wp-content/uploads/2016/06/30_government_debt_management_zlb.pdf (visited on 04/13/2026).

Hamilton, James D. (Mar. 1989). “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle”. In: *Econometrica* 57.2, pp. 357–384.

Høyland, Kjetil and Stein W. Wallace (2001). “Generating Scenario Trees for Multistage Decision Problems”. In: *Management Science* 47.2, pp. 295–307. DOI: 10.1287/mnsc.47.2.295.9834.

International Monetary Fund and World Bank (Apr. 2014). *Revised Guidelines for Public Debt Management*. IMF Policy Paper. International Monetary Fund. URL: https://www.imf.org/external/np/pp/eng/2014/040114.pdf (visited on 04/13/2026).

Itkonen, Juha and Petteri Juvonen (Dec. 18, 2017). *Nowcasting the Finnish economy with a large Bayesian vector autoregressive model*. BoF Economics Review 6/2017. Bank of Finland. URL: https://urn.fi/URN:NBN:fi:bof-201712181708 (visited on 04/15/2026).

Kim, Chang-Jin and Charles R. Nelson (1999). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. MIT Press. ISBN: 9780262112383.

Lee, Kang-Soek and Richard A. Werner (2018). “Reconsidering Monetary Policy: An Empirical Examination of the Relationship Between Interest Rates and Nominal GDP Growth in the U.S., U.K., Germany and Japan”. In: *Ecological Economics* 146, pp. 26–34. ISSN: 0921-8009. DOI: https://doi.org/10.1016/j.ecolecon.2017.08.013. URL: https://www.sciencedirect.com/science/article/pii/S0921800916307510.

Lütkepohl, Helmut (2005). *New Perspectives in VAR Modeling*. New York: Springer.

Missale, Alessandro (1999). *Public Debt Management*. Oxford University Press. ISBN: 9780198290858.

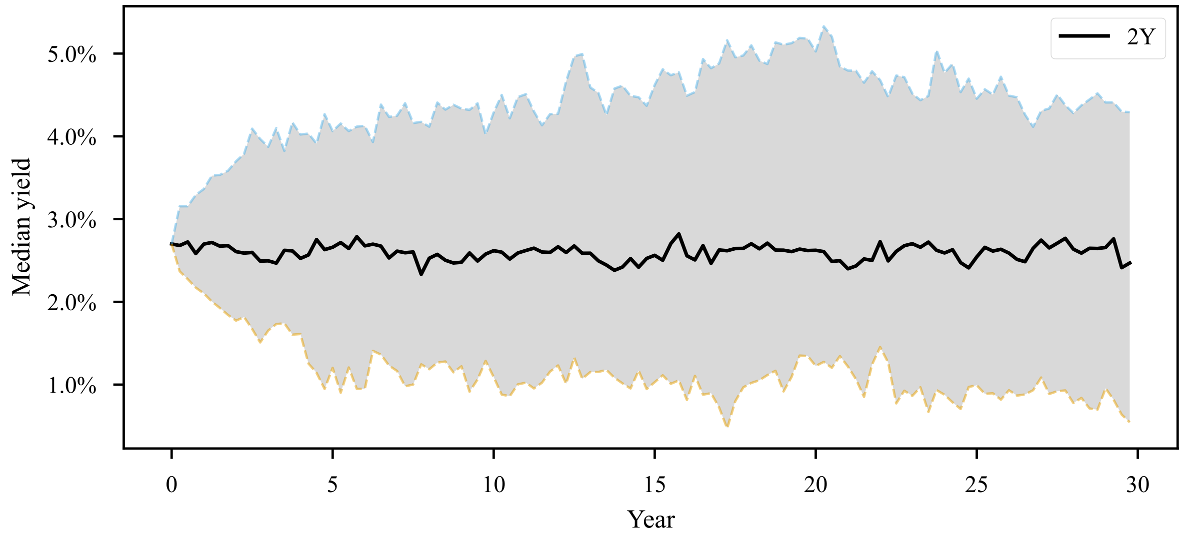
Rockafellar, R Tyrrell, Stanislav Uryasev, et al. (2000). “Optimization of Conditional Value-at-Risk”. In: *Journal of Risk* 2, pp. 21–42.

Shreve, Steven E. (2004). *Stochastic Calculus for Finance. 2, Continuous-Time Models*. eng. Springer finance. New York: Springer Vlg. ISBN: 0-387-40101-6.

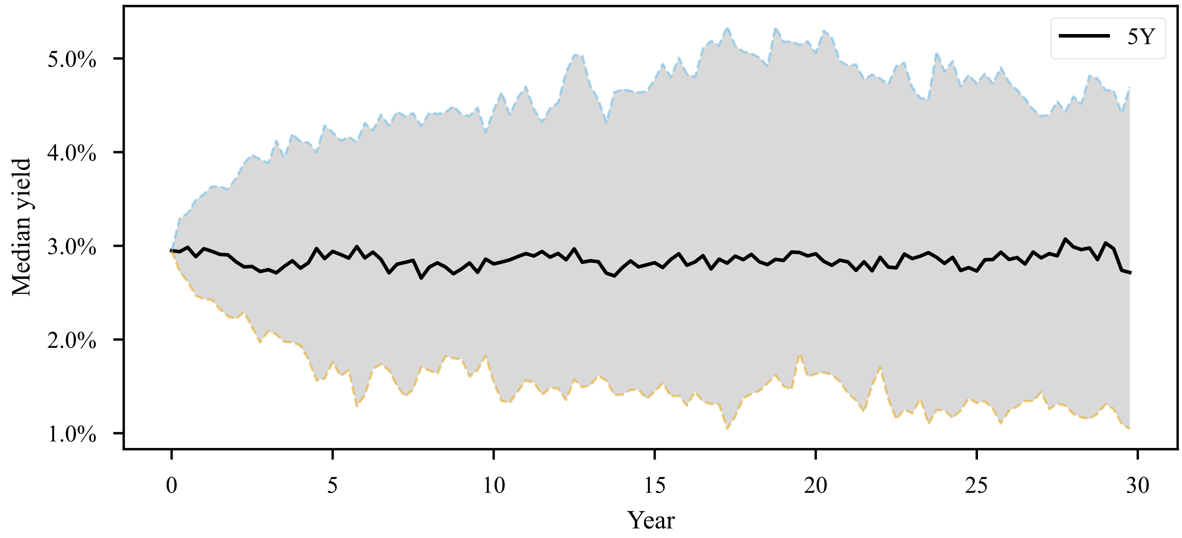
Smets, Frank and Raf Wouters (2003). “An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area”. In: *Journal of the European Economic Association* 1.5, pp. 1123–1175. DOI: 10.1162/154247603770383415.

Zenios, Stavros A., Andrea Consiglio, Marialena Athanasopoulou, Edmund Moshammer, Angel Gavilan, and Aitor Erce (2021). “Risk Management for Sustainable Sovereign Debt Financing”. In: *Operations Research* 69.3, pp. 755–773. DOI: 10.1287/opre.2020.2055. eprint: https://doi.org/10.1287/opre.2020.2055. URL: https://doi.org/10.1287/opre.2020.2055.

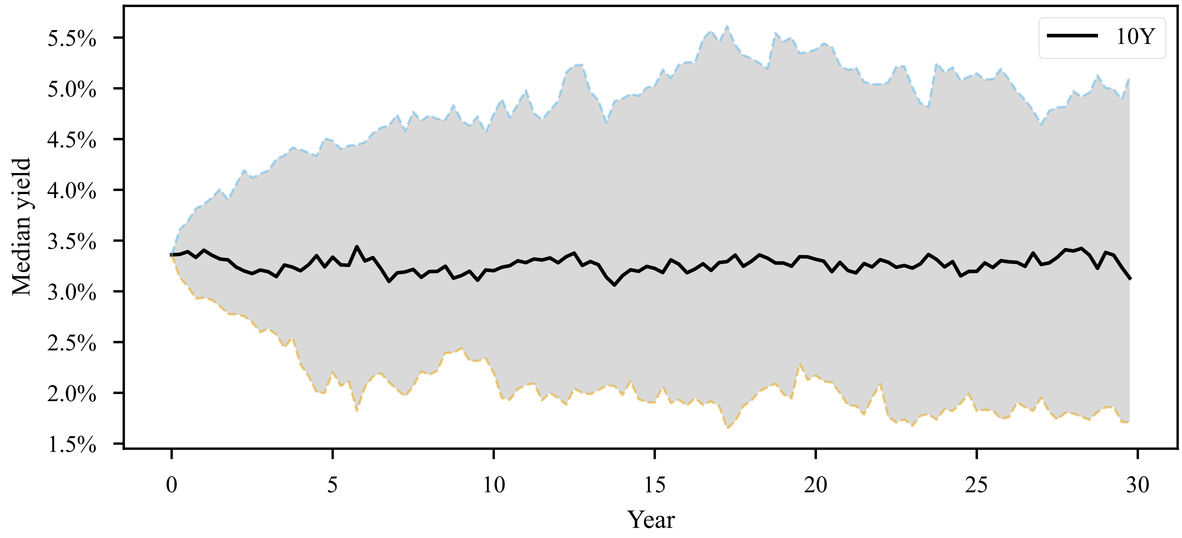
# A Additional Figures and Tables




**(a)** 2-year maturity

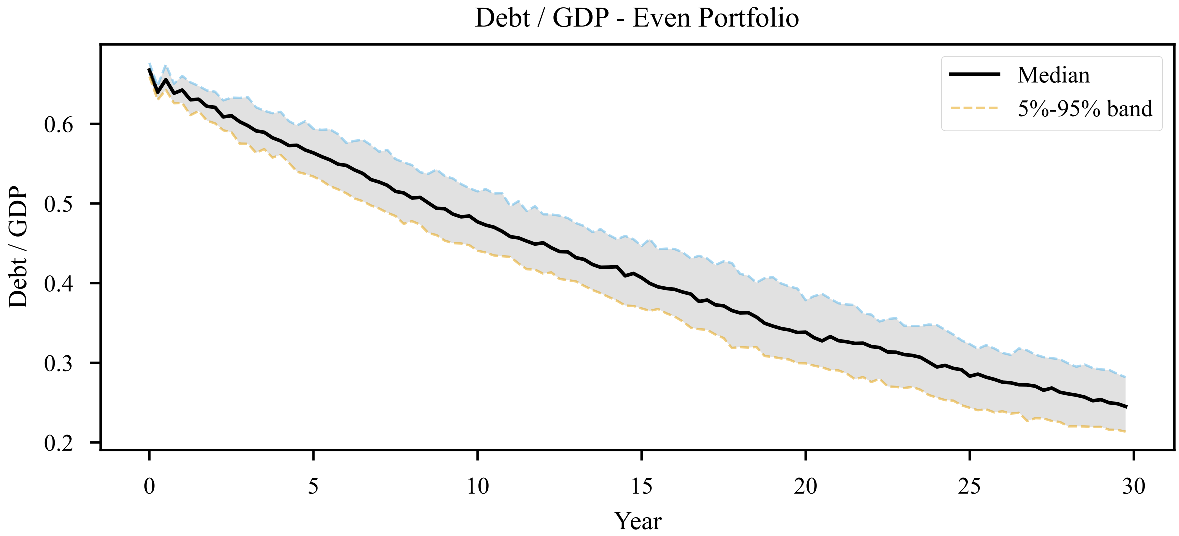



**(b)** 5-year maturity

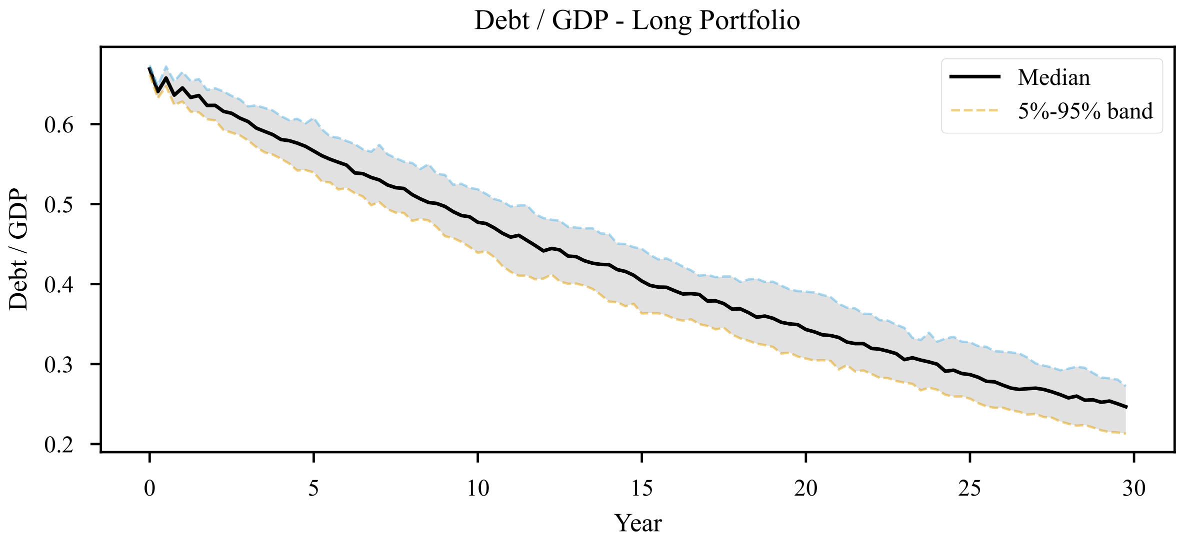



**(c)** 10-year maturity

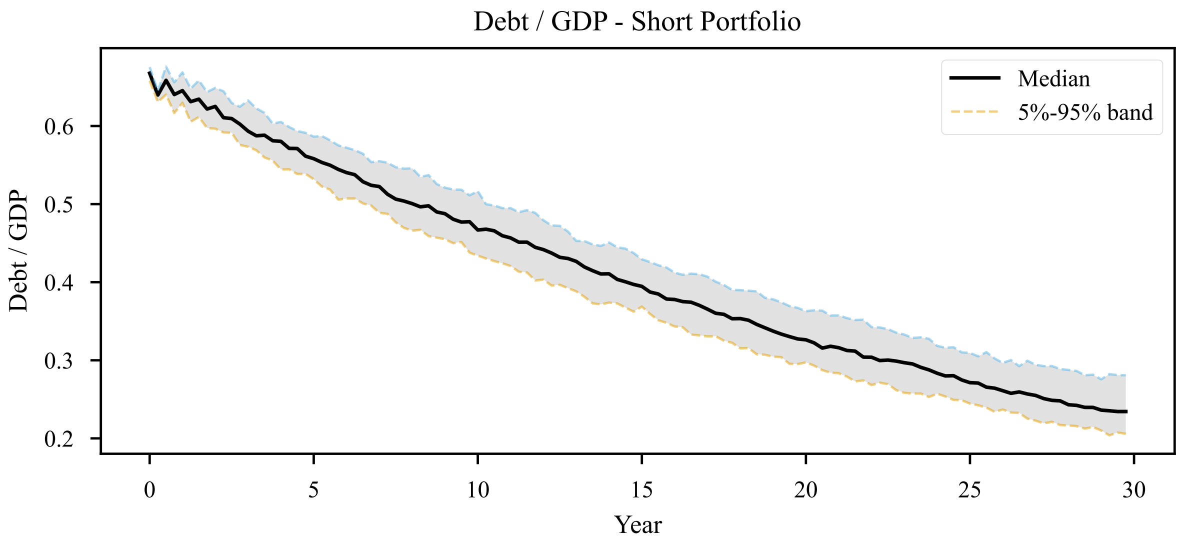
**Figure 7:** Yield forecast comparisons for 2-, 5-, and 10-year maturities with 5%–95% percentile bands.




**(a)** Even Portfolio Strategy

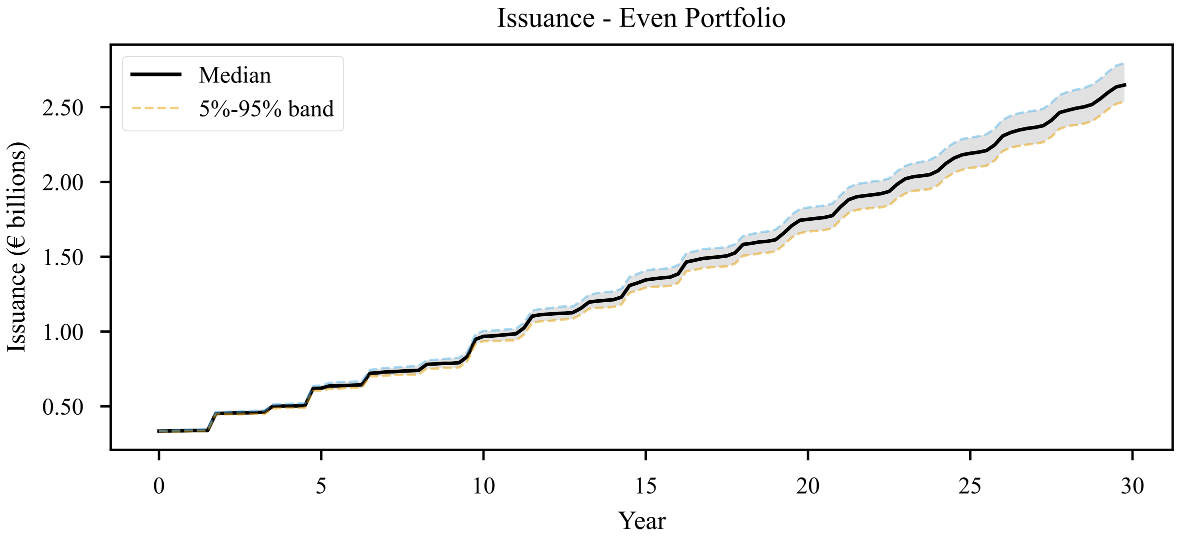



**(b)** Long Portfolio Strategy

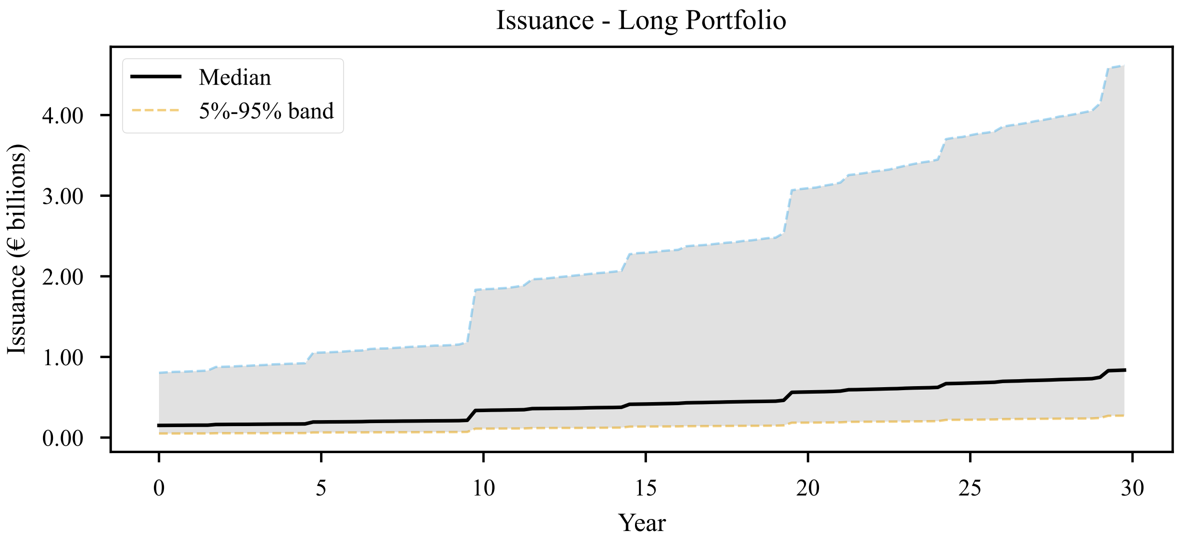



**(c)** Short Portfolio Strategy

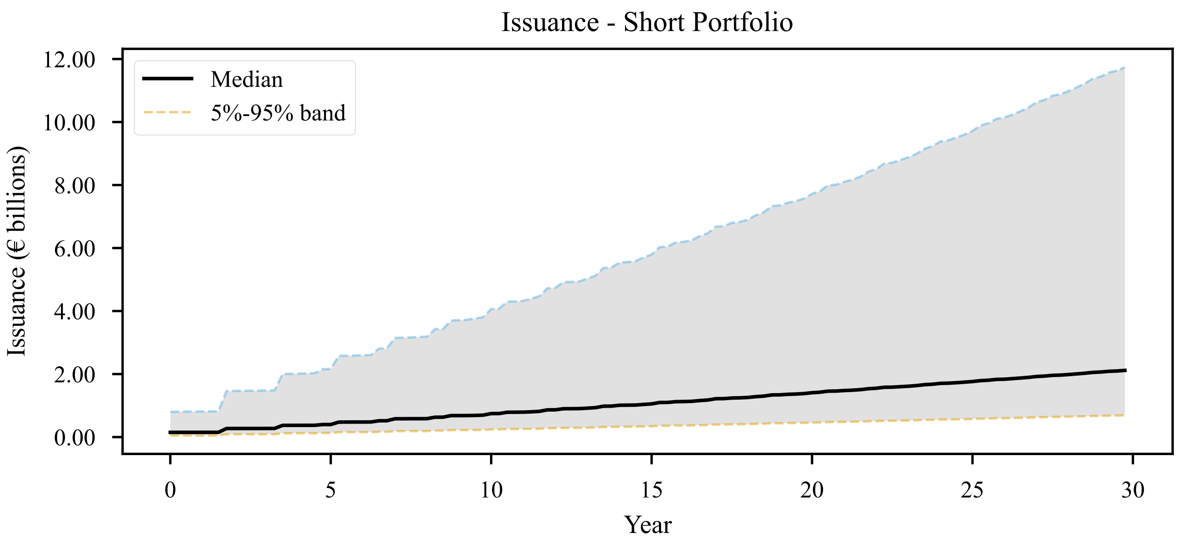
**Figure 8:** Comparison of Debt-to-GDP ratios across different portfolio structures.




**(a)** Even Portfolio Strategy

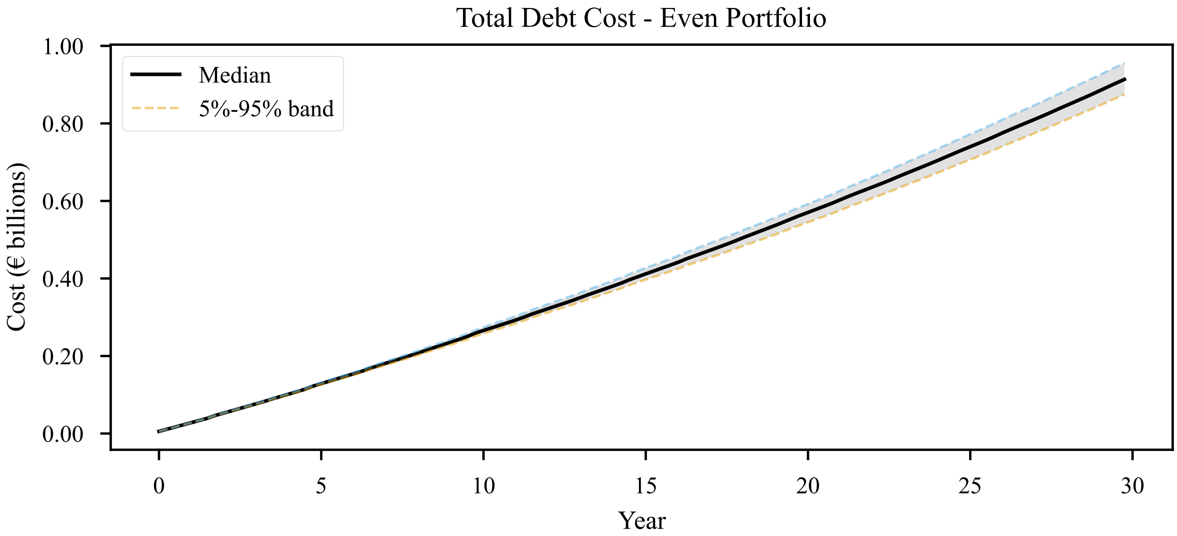



**(b)** Long Portfolio Strategy

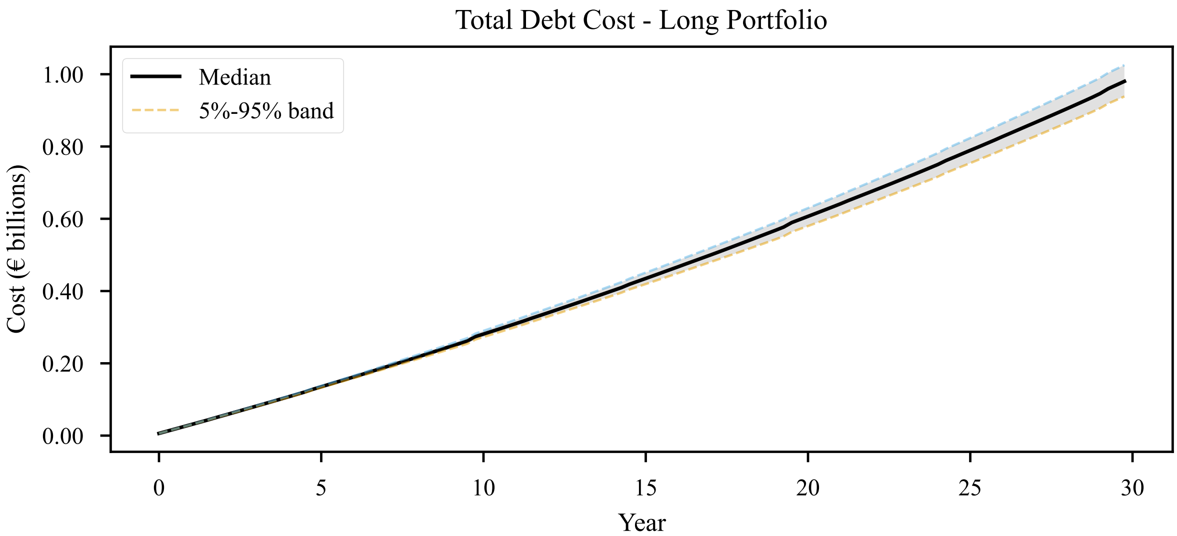



**(c)** Short Portfolio Strategy

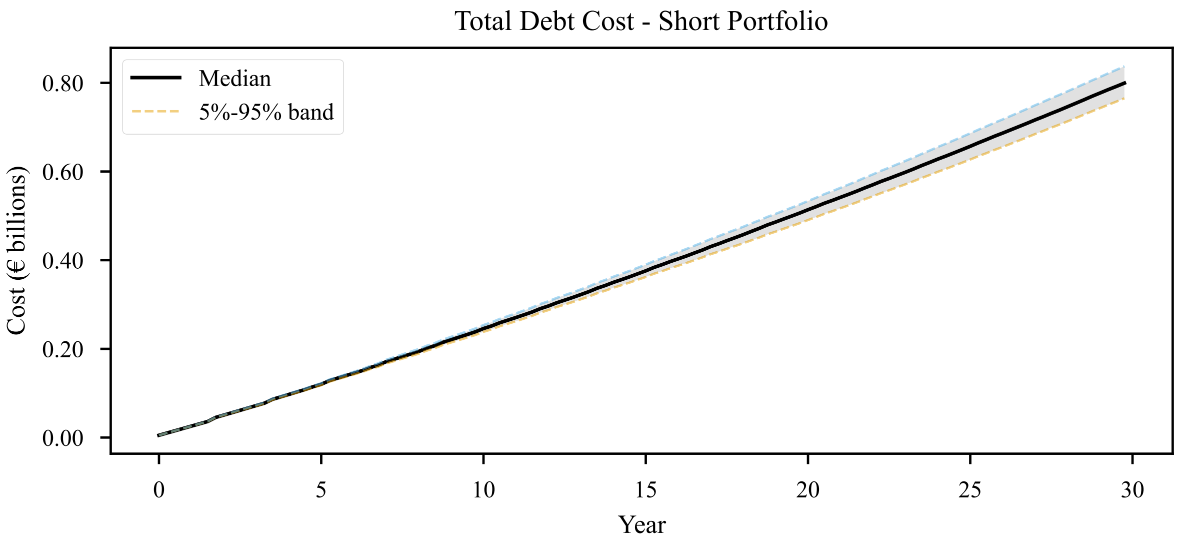
**Figure 9:** Issuance patterns with confidence bands for the evaluated portfolios.




**(a)** Even Portfolio Strategy




**(b)** Long Portfolio Strategy




**(c)** Short Portfolio Strategy

**Figure 10:** Total debt cost analysis across even, long, and short portfolio durations.

# B Self Assessment

## B.1 Implementation vs. the initial project plan

The project followed the initial project plan at the architectural level: the work was organized around a modular pipeline in which (i) macro-financial scenarios are generated, (ii) scenarios are mapped into debt cash flows and stock dynamics, and (iii) strategies are compared using cost and risk metrics with supporting visualizations. The staged approach (framing → modeling → implementation → results and reporting) remained consistent throughout.

## B.2 In what regard was the project successful?

The project was successful in delivering a coherent and operational decision-support prototype. The main success is that the complete pipeline runs end-to-end and produces interpretable outputs for strategy comparison. The model can generate macro-financial scenarios, evaluate alternative portfolio and issuance policies, and report the key decision metrics (interest cost outcomes across horizons, Debt/GDP dynamics, and GFN/GDP profiles) in a form that supports discussion of trade-offs.

Finally, communication with the client supported progress. Weekly interaction ensured that the work remained oriented toward decision-relevant outputs and that intermediate results could be challenged and interpreted throughout development rather than only at the end.

## B.3 In what regard was the project less successful

The primary limitations relate to data relevance and the realism of macroeconomic scenario generation. While the framework is operational, the credibility of quantitative outcomes depends on the quality and representativeness of inputs. Data constraints and uncertainty about the appropriate mapping between available series and the required model variables limited calibration depth and reduced confidence in the level of the reported metrics.

A second limitation is that the GDP scenario block appears to produce outcomes that are optimistic relative to what a risk-focused sovereign debt analysis would ideally incorporate. This is not necessarily a coding issue; it reflects a more general challenge in macroeconomic modeling. GDP growth is influenced by structural factors, policy choices, and shocks that are difficult to represent parsimoniously. Regime-switching models impose a simplified structure that can understate downside tail risk if regimes, transition probabilities, or shock distributions are not calibrated to capture rare but policy-relevant adverse events. As a result, sustainability outcomes should be interpreted conditionally on the macro assumptions rather than as definitive forecasts.

## B.4 What could have been done better

### B.4.1 Project team

In hindsight, the team could have formalized integration interfaces earlier. A stricter specification of variable definitions and timing conventions for example, the precise definition of refinancing rate, the timing of cash flows, and the mapping between GDP paths and debt-to-GDP reporting would likely have reduced integration overhead. More structured integration testing for example, automated unit tests for accounting identities and consistency checks between modules would further strengthen reliability. The team could also have devoted additional effort to macro scenario realism. Given the sensitivity of sustainability metrics to GDP dynamics, a dedicated robustness would have been valuable.

#### B.4.2 Client

Client interaction was frequent, but earlier agreement on a minimal calibration and validation target set could have improved focus. In particular, agreeing on a small number of stylized facts and benchmarks (for example, plausible ranges for growth regimes, interest–growth differentials, and refinancing pressure metrics) would help anchor the scenario engines and reduce ambiguity about what constitutes a realistic scenario set.

#### B.4.3 Teaching staff

We did not actively seek support from the teaching staff during the implementation phase, and we met the course teacher only once. As a result, the project was largely executed based on our own interpretation of the course requirements. This approach had the advantage of keeping the work tightly focused on delivering an operational prototype in relatively efficient time.