BETA
Bureau d'économie théorique et appliquée

Documents de travail

«Using DSGE and Machine Learning to Forecast Public Debt for France»

Auteurs

Emmanouil SOFIANOS, Thierry BETTI, Theophilos PAPADIMITRIOU, Amélie BARBIER-GAUCHARD, Periklis GOGAS

Document de Travail n° 2025 – 18

Avril 2025

Bureau d'Économie Théorique et Appliquée BETA

https://www.beta-economics.fr/

Contact :
jaoulgrammare@beta-cnrs.unistra.fr

Using DSGE and Machine Learning to Forecast Public Debt for France

Emmanouil SOFIANOS *
Thierry BETTI †
Theophilos PAPADIMITRIOU ‡
Amélie BARBIER-GAUCHARD §
Periklis GOGAS ¶

April 16, 2025

Abstract: Forecasting public debt is essential for effective policymaking and economic stability, yet traditional approaches face challenges due to data scarcity. While machine learning (ML) has demonstrated success in financial forecasting, its application to macroeconomic forecasting remains underexplored, hindered by short historical time series and low-frequency (e.g., quarterly/annual) data availability. This study proposes a novel hybrid framework integrating Dynamic Stochastic General Equilibrium (DSGE) modeling with ML techniques to address these limitations, focusing on the evolution of France's public debt. We first generate a large synthetic macroeconomic dataset using an estimated DSGE model for France, which allows for efficient training of ML algorithms. These trained models are then applied to actual historical data for directional debt forecasting. The results show that the best machine learning model is an XGBoost achieving 90% accuracy. Our results highlight the viability of combining structural economic models with data-driven techniques to improve macroeconomic forecasting.

Keywords: DSGE, Machine Learning, Public Debt, Forecasting, France.

Classification JEL: C53, E27, E37 H63, H68

*BETA, University of Strasbourg, University of Lorraine, CNRS, 67000, Strasbourg, (France), sofianos@unistra.fr.

†BETA, University of Strasbourg, University of Lorraine, CNRS, 67000, Strasbourg, (France), tbetti@unistra.fr.

‡Department of Economics, Democritus University of Thrace (Greece), papadimi@econ.duth.gr.

§BETA, University of Strasbourg, University of Lorraine, CNRS, 67000, Strasbourg, (France), abarbier@unistra.fr.

¶Department of Economics, Democritus University of Thrace (Greece), pgkogkas@econ.duth.gr.

1 Introduction

The crises that have impacted developed economies since the Great Recession of 2008 - including the COVID-19 pandemic, the Russia-Ukraine war, and escalating climate disasters - have required significant government interventions. These measures, aimed at stabilizing economies and mitigating social hardships, have led to considerable increases in public debt. In the Economic and Monetary Union (EMU), the situation of public finance remains a burning issue, especially due to the increasing heterogeneity of fiscal positions across Member States. The European fiscal rule, introduced when the Eurozone was created in 1999, was suspended from 2020 to 2023 due to the escape clause. This fiscal rule, introduced by the Stability and Growth Pact (SGP, 1996), must combine the sound management of national public finance on the one hand, but also leave Member States sufficiently margin to achieve their fiscal policy on the other. This dual objective, particularly tricky to achieve, was reaffirmed by the reform adopted in 2024 and which came into force on April 30, 2024. In the reformed fiscal rule, a particular focus is paid to the sustainability of public debt in the medium term. The *national medium-term fiscal-structural plans* and the *medium-term net primary expenditure path* are now the new public finance monitoring tools anchored on public debt sustainability.

Although many European countries are showing an increasing trend towards public debt, the situation is even more alarming in France. France ranks third, after Greece and Spain, among the countries in the Eurozone with the highest level of public debt. In addition, the French political context has been extremely uncertain since the legislative elections of June 2024. For these reasons, the case study proposed in this article focuses on France. Existing research on public debt determinants emphasizes that socio-economic factors do not solely determine public debt dynamics; political and institutional factors also play a crucial role, as demonstrated in studies by Di Bartolomeo et al. (2018) and Barbier-Gauchard & Sofianos (2024). According to Estefania-Flores et al. (2023), there is a significant positive forecast error in the projections of the debt-to-GDP ratio in both advanced economies and emerging and developing economies. This error in public debt forecasts, which tends to increase during periods of election, can have serious repercussions. This may be the result of governments choosing to follow different regimes that do not pursue fiscal consolidation during election periods. High and unpredicted levels of public debt often lead to high interest rates on government bonds, increasing the cost of borrowing for governments, which increases the debt burden and, consequently, further increases the level of public debt. This in turn can limit their ability to finance future projects. It can also lead to low investor confidence, as investors become increasingly concerned about the government's ability to repay its debt, which can lead to reduced foreign investment and economic instability. All these factors can have a considerable influence on the way that the financial markets assess a country's exposure to the default risk of default, and the rating agencies that rate that risk.

In these conditions, being able to accurately forecast public debt seems to be a challenge. An accurate methodology could help policymakers and international stakeholders assess fiscal health and take proactive measures to address potential vulnerabilities in public debt. The IMF is the leading authority for global public debt forecasting (IMF (2024)). In the EU, the European Commission also proposes its own public debt forecasts (European Commission (2024)). Unfortunately, forecasting errors continue to exist, we are far from having accurate public debt forecasting. Bachleitner & Prammer (2024) underline that the error in the public debt forecast is more pronounced for high-debt countries than for low-debt countries. On the other hand, the use of machine learning to provide forecasts is growing more and more. Machine learning is gaining more and more attention in economics. Recent applications include forecasting macroeconomic variables, creation of early warning systems for financial crises, recessions, or risk of default. Historically, large data sets were necessary to efficiently train machine learning models (Gogas & Papadimitriou (2021)). The availability of high-frequency data facilitated the application of machine learning techniques in finance. However, during the past few years, machine learning techniques have been used successfully with shorter macroeconomic datasets, providing results that outperform other econometric techniques (Sermpinis et al. (2014) for inflation and unemployment, Gogas et al. (2021) for output gaps, Gogas et al. (2022) for unemployment, Lekhal (2024) for public external debt, Silva et al. (2024) for international trade, or Belly et al. (2023) for sovereign risks.

Machine learning models are particularly well suited for early warning predictions in economics, including financial crises, recessions, and default risks. These models excel in handling complex high-dimensional datasets, often outperforming traditional econometric techniques in such tasks. By automating model building, machine learning algorithms can learn from historical data, identify patterns, and make decisions with minimal human intervention. Unlike traditional approaches, ML-based methodologies are data-driven and largely atheoretical, extracting insights directly from the data without relying on predefined theoretical frameworks. Furthermore, machine learning's ability to model non-linear relationships and adapt to evolving patterns makes it an invaluable tool in economics, where predictive accuracy, flexibility, and adaptability are crucial for addressing complex and dynamic challenges.

The small size of the macroeconomic databases remains a major drawback to overcome. DSGE models seem to be the ideal methodological tool to address this difficulty, as they provide a simulation of the functioning of the real economy and, thus, can generate large amounts of data. To our knowledge, the combination of the DSGE framework and machine learning has not yet been widely used. Hinterlang & Hollmayr (2022) is the first paper that applies machine learning techniques to classify an unobserved economic state using DSGE generated data. The authors identify US monetary and fiscal dominance regimes using machine learning techniques. The algorithms are trained and verified using simulated data from Markov-switching DSGE models, before they are used to classify regimes from 1968–2017 on actual US data. All machine learning methods outperformed the standard logistic regression with respect to the simulated data. Stempel & Zahner (2023) couples a new-Keynesian model with a neural network to assess whether the European Central Bank (ECB) conducted monetary policy between 2002 and 2022 according to the weighted average of inflation rates within the EMU or reacted more strongly to developments in inflation rates in certain countries.

This paper stands at the crossroads of three fields of literature: the empirical literature on public debt forecasting, the theoretical literature that integrates government public debt into DSGE models (Corsetti et al. (2012) for instance), and the literature that applies machine learning techniques in macroeconomics. The aim of this paper is to propose a ML public debt forecasting model for France trained on artificial data produced by the DGSE model as illustrated by Figure (1).

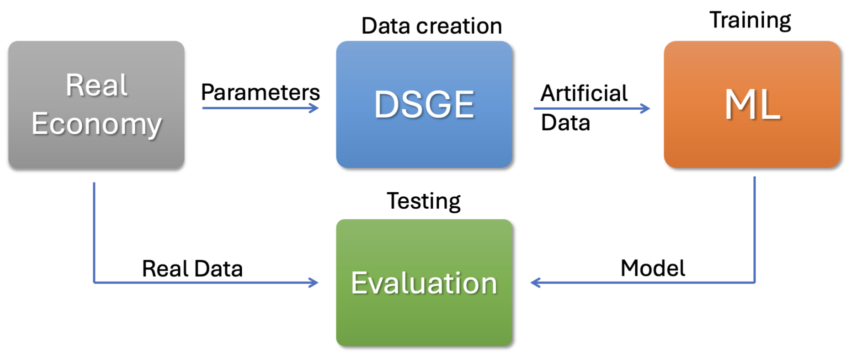

Figure 1: Combination of the DSGE model and Machine Learning

The rest of the paper is organized as follows. Section (2) presents the DSGE model, its estimation, and the synthetic data to train the machine learning models. Section (3) describes the use of machine learning techniques and main results, and section (4) concludes.

2 The DSGE model and simulations

This section presents the baseline DSGE model, the estimation of key parameters for France used in the calibration to simulate the model and the results of simulations which provide synthetic data to train the machine learning models, in order to provide better knowledge of the evolution of future public debt.

2.1 The log-linear model

We use a standard DSGE model close to the canonical model developed in Smets & Wouters (2007). Since we aim to forecast government debt, we introduce the dynamic of the public debt and a risk premium on the interest rate of the government bond. In this section, we describe the set of log-linear equations used to simulate the model. All variables are log-linearized around their steady-state balanced growth path. Star variables denote steady-state values. We first describe the aggregate demand side of the model (consumption, investment, public spending), then turn to the aggregate supply and the economic policies (fiscal and monetary). The canonical model of Smets & Wouters (2007) has demonstrated strong forecasting performance. While it would be possible to develop a more sophisticated model - for instance, by incorporating a more developed financial sector (see, e.g., Gertler & Karadi (2011)) - it is not clear that such an extension would necessarily improve predictive accuracy (see Kolasa & Rubaszek (2014) for a discussion). Similarly, we could consider an open-economy framework, but we opt here for a tractable and easily estimable model. For more details about the microfoundations of such a standard DSGE model, see Smets & Wouters (2007).

Households in this economy maximize utility over consumption and labor supply, subject to an intertemporal budget constraint. They also provide capital to firms through lending. Equation (1) represents the Euler equation which describes the dynamic of households consumption as:

$$c_t = c_1\, c_{t-1} - (1-c_1)\mathbb{E}_t c_{t+1} - c_2\,(l_t - E_t l_{t+1}) - c_3(r_t - E_t\pi_{t+1} + \xi_t^b), \tag{1}$$

with

$$c_1 = \frac{\lambda}{1+\lambda}, \quad c_2 = \frac{(\sigma_c - 1)(W^* L^*/C^*)}{\sigma_c(1+\lambda)}, \quad c_3 = \frac{1-\lambda}{(1+\lambda)\sigma_c}.$$

Thus, the level of consumption c_t depends on the expected level of consumption, but we also introduce a habit formation mechanism, which implies that current consumption also depends on its past value. l_t are worked hours, r_t the nominal (riskless) interest rate, π_t the consumer price index inflation rate. c_1 and c_2 contains: λ measures the degree of habit formation and σ_c the elasticity of intertemporal substitution. W^*, L^* and C^*, are the values of, respectively, wages, hours worked and consumption at the steady-state. Finally, ξ_t^b is an exogenous process that affects the wedge between the interest rate controlled by the central bank and the return on riskless assets held by households. The exogenous shock to household consumption ξ_t^b follows an AR(1) process such as:

$$\xi_t^b = \rho^b \xi_{t-1}^b + \epsilon_t^b \tag{2}$$

with $\rho^b \in (0,1)$ which capture the duration of the shock and ϵ_t^b an i.i.d exogenous disturbance.

Thus, equation (3) defines the dynamic of private investment i_t:

$$i_t = i_1\, i_{t-1} + (1-i_1)\mathbb{E}_t i_{t+1} + i_2\, q_t + \xi_t^i, \tag{3}$$

where

$$i_1 = \frac{1}{1+\beta}, \quad i_2 = \frac{1}{1+\beta\psi}.$$

with β is the standard discount factor applied by households, ψ is the steady-state elasticity of the capital adjustment cost function, q_t is the Tobin Q and ξ_t^i is a disturbance to the investment-specific technology process, such as:

$$\xi_t^i = \rho^i \xi_{t-1}^i + \epsilon_t^i \tag{4}$$

with $\rho^i \in (0,1)$ which capture the duration of the shock and ϵ_t^i an i.i.d exogenous disturbance.

The Tobin's Q, representing the evolution of the value of the capital, is defined as:

$$q_t = q_1\, \mathbb{E}_t q_{t+1} + (1-q_1)\mathbb{E}_t r_{t-1}^k - (r_t - E_t\pi_{t+1} + \xi_t^b), \tag{5}$$

with

$$q_1 = \beta(1-\delta)$$

r_t^k is the real rental rate on capital. Then, the value of the capital depends on the difference between the real rental rate on capital and the riskless real interest rate. The shock ξ_t^b given by equation (2), which impacts the return on assets held by households, will also influence the investment decision. The parameter δ represents capital depreciation.

The third component of the aggregate demand, government expenditure, is simply introduced as an AR(1) process such as:

$$g_t = \rho^g g_{t-1} + \epsilon_t^g \tag{6}$$

with $\rho^g \in (0,1)$ which capture the duration of the shock and ϵ_t^g an i.i.d exogenous disturbance.

Supply side

The production function of firms is given by a traditional Cobb-Douglas function combining labor l_t and capital services used in production $k_{s,t}$, such as:

$$y_t = \alpha k_{s,t} + (1-\alpha) l_t + \xi_t^a, \tag{7}$$

with α is the productivity parameter, ξ_t^a a total productivity shock which impacts the marginal productivity of inputs defined as:

$$\xi_t^a = \rho^a \xi_{t-1}^a + \epsilon_t^a \tag{8}$$

with $\rho^a \in (0,1)$ which capture the duration of the shock and ϵ_t^a an i.i.d exogenous disturbance.

Current capital services used in production are a function of capital installed in the previous period k_{t-1} and the degree of capital utilization z_t, such as:

$$k_{s,t} = k_{t-1}\, z_t \tag{9}$$

The degree of capital utilization, is function of the real rental rate of capital r_t^k as :

$$z_t = \frac{1-\psi}{\psi}\, r_t^k \tag{10}$$

with ψ is a parameter in the function which describes the adjustment cost related to the changes in the degree of capital utilization.

The law of mention of the capital stock is:

$$k_t = (1-\delta) k_{t-1} + \delta i_t + \delta(1+\beta)\psi \xi_t^i \tag{11}$$

with ξ_t^i the disturbance to the investment-specific technology process given by equation (4). Cost minimization by firms will also imply that the rental rate of capital is negatively related to the capital-labor ratio and positively to the real wage (both with unitary elasticity). Thus, the optimal input choice follows:

$$r_t^k = -(k_t - l_t) + w_t \tag{12}$$

Price setting

Firms evolve in a monopolistic competitive goods market. They are price-setters but face a degree of nominal rigidity *à la* Calvo (1983). Following the optimization by firms and aggreggation, the law of motion of prices is given by:

$$\pi_t = \pi_1\, \pi_{t-1} + \pi_2\, \mathbb{E}_t \pi_{t+1} - \pi_3\, \mu^p_t + \xi^p_t \tag{13}$$

with

$$\pi_1 = \frac{\theta}{1+\beta\theta}, \pi_2 = \frac{\beta}{1+\beta\theta}, \pi_3 = \frac{1}{1+\beta\theta}\frac{(1-\beta\kappa)(1-\kappa)}{\kappa}.$$

Inflation, denoted by π_t, is defined as the change in the consumer price index, that is, $\pi_t = p_t - p_{t-1}$, where p_t is the (log) consumer price index. θ is the degree of indexation to past inflation and κ the degree of price rigidity (in the price-setting *à la* Calvo (1983).

with coefficients π_1, π_2, and π_3 determined by the degree of price rigidity and indexation and ξ^p_t is a cost-push shock, defined as a standard AR(1) process, such as:

$$\xi^p_t = \rho^p \xi^p_{t-1} + \epsilon^p_t \tag{14}$$

with $\rho^p \in (0,1)$ which capture the duration of the shock and ϵ^p_t an i.i.d exogenous disturbance. In addition, μ^p_t defines the price mark-up given by the difference between the marginal product of labor and the real wage such as:

$$\mu^p_t = \alpha(k_{s,t} - l_t) - w_t + \xi^a_t \tag{15}$$

with ξ^a_t, the total productivity shock which impacts the marginal productivity of inputs defined by equation (8).

Wage setting

On the labor market, workers are also in a monopolistic competition environment then wage-setters. Similarly to firms, workers face a degree of nominal wage rigidity introduced *à la* Calvo (1983). The resulting wage dynamic is given by:

$$w_t = w_1\, w_{t-1} + (1-w_1)(\mathbb{E}_t w_{t+1} + \mathbb{E}_t \pi_{t+1}) - w_2 \pi_t + w_3 \pi_{t-1} - w_4 \mu^w_t + \xi^w_t \tag{16}$$

with

$$w_1 = \frac{1}{1+\beta}, w_2 = \frac{1+\beta\theta^w}{1+\beta}, w_3 = \frac{\theta^w}{1+\beta}, w_4 = \frac{1}{1+\beta}\frac{(1-\beta\kappa^w)(1-\kappa^w)}{\kappa^w}.$$

Similarly to the setting of prices, θ^w is the degree of indexation to past inflation and κ^w the degree of wage rigidity. Analogously to price-setting, μ^w_t is the wage markup, defines as the real wage and the marginal rate of substitution between work and consumption, such as:

$$\mu^w_t = w_t - \left(\sigma_l l_t + \frac{1}{1-\lambda}(c_t - \lambda c_{t-1})\right) \tag{17}$$

with σ_l the elasticity of labor supply with respect to the real wage. Finally, ξ^w_t is an exogenous disturbance on wages, defined such as:

$$\xi^w_t = \rho^w \xi^w_{t-1} + \epsilon^w_t \tag{18}$$

with $\rho^w \in (0,1)$ which capture the duration of the shock and ϵ^w_t an i.i.d exogenous disturbance.

Monetary policy

The central bank sets the nominal interest rate in the economy, following a standard Taylor rule given by equation (19) :

$$r_t = \gamma^r \, r_{t-1} + \gamma^\pi \, (\pi_t - \pi^*) + \gamma^y \, (y_t - y^*) + \xi^r_t, \tag{19}$$

where r_t is the nominal interest rate, γ_r measures the persistence of the nominal interest rate, γ^π and γ^y measures the sensibility of the central bank for the stabilization of, respectively, inflation and the output gap. ξ^r_t is a monetary policy shock defined as an AR(1) process such as:

$$\xi^r_t = \rho^r \xi^r_{t-1} + \epsilon^r_t \tag{20}$$

with $\rho^r \in (0,1)$ which capture the duration of the shock and ϵ^r_t an i.i.d exogenous disturbance.

Fiscal policy

The government finances public spending defined by equation (6) through public debt. The dynamic of public debt is given by :

$$b_t = r^b_{t-1} + \left(\frac{1}{\beta}\right) \left[b_{t-1} - \pi_{t-1} + \left(\frac{G^*}{Y^*}\right) g_t \right] \tag{21}$$

The public debt b_t is given by the lagged public debt and the current level of public spending, β represents the discount factor. r^b_t denotes the return on government securities which depends on the nominal interest rate sets by the central bank, but also on a risk premium that endogenously increases with the stock of public debt, such as:

$$r^b_t = r_t + \Delta \, (b_t - p_t - y_t) \tag{22}$$

where Δ measures the sensibility of the risk premium on government bonds to the evolution of public debt.

Aggregate variables and market clearing conditions

The aggregate resource constraint (market clearing condition) is given by:

$$y_t = \frac{C^*}{Y^*} c_t + \frac{I^*}{Y^*} i_t + z_y z_t + \left(1 - \frac{C^*}{Y^*} - \frac{I^*}{Y^*}\right) g_t, \tag{23}$$

where output y_t equals aggreggate demande denided as the sum of consumption (c_t), investment (i_t), capital-utilization costs (a function of z_t), and public spending (g_t). $\frac{C^*}{Y^*}$ is the steady-state share of consumption in output, $\frac{I^*}{Y^*}$ the investment-output ratio and $\frac{G^*}{Y^*}$ the government expenditure-output ratio. Through calculation, we can define $z_y = R^*_k \frac{K^*}{Y^*}$, with R^*_k the steady-state rental rate of capital and $\frac{K^*}{Y^*}$ the capital-output ration at the steady-state (see Smets & Wouters (2007)).

2.2 Estimation procedure

We estimate the model using quarterly data for the French economy from 2000Q1 to 2023Q2. The Bayesian estimation method follows standard practice in the literature (see for instance An & Schorfheide (2007)). The model is estimated using eight macroeconomic time series: GDP, household consumption, investment, hours worked, inflation, the nominal (riskless) interest rate, and public debt. Table (1) displays the source of the original data. All variables are expressed in logarithms. Nominal series for GDP, household consumption, investment, public debt, and compensation to employees are deflated using the GDP deflator. With the exception of the nominal interest rate, we apply the Hamilton filter to extract the cyclical component of each series (see Hamilton (2017) for a description of the filter). Inflation is the quarterly growth rate of the GDP deflator. Finally, real wages are constructed by deflating compensation to employees and dividing by the total labor force.

Variable	Description	Source
y	Nominal GDP, Seas. and cal. adjusted)	Eurostat database
r^b	10 years, annual rate, French rate	OECD database
b	Central government consolidated gross debt, in real terms (GDP deflator)	Eurostat
c	Final consumption by households (GDP deflator, millions euros, Seas. And cal. adj.)	Eurostat, ESA 2010
π	GDP deflator	Eurostat
r	ECB policy rate	ECB database
l	Total hours worked	ECB database
i	Gross fixed capital formation (GDP deflator, millions euros, Seas. and cal. adjusted)	Eurostat, ESA 2010
w	Own calculations	OECD database

Table 1: Source of the dataset

We use prior means close to the original estimation in Smets & Wouters (2007). All model parameters are estimated, with the exception of δ, the capital depreciation rate, which is calibrated at 0.025, and $\Delta = 0.3$ in order to match empirical moments of government yields. Furthermore, we calibrate the steady-state ratios $\frac{C^*}{Y^*} = 0.52$ and $\frac{I^*}{Y^*} = 0.23$ based on the average values computed from the empirical series employed in the model estimation. Henceforth, $\frac{G^*}{Y^*} = 1 - \frac{C^*}{Y^*} - \frac{I^*}{Y^*} = 0.25$.

Tables (2) and (3) display the prior and posterior distributions for, respectively, the estimated parameters and the shock processes. In Table (3), $[\sigma^a, \sigma^b, \sigma^g, \sigma^I, \sigma^r, \sigma^p, \sigma^w]$ are the estimated standard deviations of the different exogenous shocks.

Table 2: Prior and Posterior Distribution of Structural Parameters

Parameter	**Prior Distribution** Mean	Std. Dev.	**Posterior Distribution** Mode	Mean	5%	95%
ψ	4.00	1.50	6.3103	6.3456	6.2894	6.4149
σ_c	1.50	0.37	1.0469	1.0460	1.0059	1.0644
λ	0.50	0.10	0.8710	0.6755	0.6667	0.6789
σ_l	2.00	0.75	3.3895	3.3860	3.3598	3.4066
θ_w	0.50	0.10	0.3894	0.3873	0.3830	0.3897
θ	0.50	0.10	0.3473	0.3428	0.3248	0.3643
κ_w	0.50	0.15	0.7209	0.8744	0.8680	0.8814
κ	0.50	0.15	0.7209	0.7207	0.7132	0.7292
γ^π	1.50	0.25	2.0120	2.04	1.74	2.33
γ^r	0.75	0.10	0.7713	0.81	0.77	0.85
γ^y	0.125	0.05	0.2636	0.08	0.05	0.12
$100(\beta^{-1}-1)$ (discount)	0.25	0.10	0.5635	0.16	0.07	0.26
α	0.24	0.01	0.2220	0.2228	0.2161	0.2308

Table 3: Prior and Posterior Distribution of exogenous shocks

	Distribution Type	Prior Mean	Std. Dev.	Posterior Mode	Mean	5%	95%
σ^a	InvGamma	0.001	1.00	0.0124	0.0124	0.0108	0.0133
σ^b	InvGamma	0.001	1.00	0.0135	0.0137	0.0126	0.0144
σ^g	InvGamma	0.001	1.00	0.0124	0.0078	0.0048	0.0115
σ^I	InvGamma	0.001	1.00	0.0062	0.0061	0.0059	0.0068
σ^r	InvGamma	0.001	1.00	0.0055	0.0055	0.0050	0.0061
σ^p	InvGamma	0.001	1.00	0.0070	0.0071	0.0064	0.0076
σ^w	InvGamma	0.001	1.00	0.0458	0.0452	0.0418	0.0513
ρ^a	Beta	0.50	0.20	0.97	0.9687	0.9626	0.9739
ρ^b	Beta	0.50	0.20	0.259	0.2547	0.2391	0.263
ρ^g	Beta	0.50	0.20	0.7139	0.7165	0.7082	0.7299
ρ^I	Beta	0.50	0.20	0.6072	0.6047	0.5922	0.6138
ρ^r	Beta	0.50	0.20	0.6646	0.0678	0.645	0.0734
ρ^p	Beta	0.50	0.20	0.8788	0.8786	0.8756	0.8845
ρ^w	Beta	0.50	0.20	0.6365	0.6423	0.6068	0.6739

2.3 Simulation of the data

To generate the synthetic dataset used for training the Machine Learning models, we simulate the estimated DSGE model over 11,000 periods, discarding the initial 1,000 observations to eliminate the influence of initial conditions. The model is subjected to a sequence of stochastic shocks across all dimensions specified within its structure. For illustrative purposes, Figure (2) presents the simulated trajectory of public debt over the first 5,000 periods.

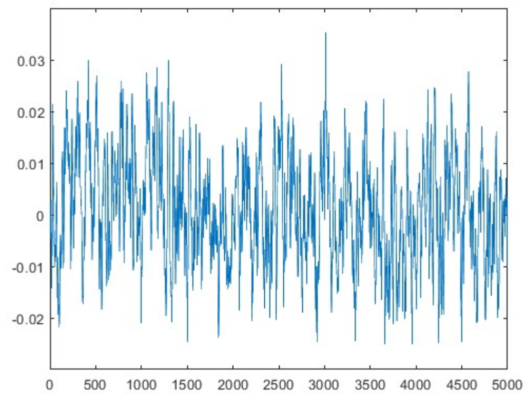

Figure 2: Simulated data for the public debt over 5000 periods

3 Machine Learning

3.1 Methodology

Machine learning, a strain of theoretical framework and methodological tools that first appeared in the 1950s, was designed to enhance the ability of artificial intelligence (AI) systems to learn autonomously. The essence of machine learning lies in its ability to automatically construct models, enabling computers to learn from historical data, identify complex and non-linear patterns, and make decisions with minimal human intervention using AI Gogas & Papadimitriou (2021). These models are flexible and capable of performing both classification and regression tasks, making them invaluable in a wide range of applications. In our research, we employ four machine learning algorithms for their comparative advantages while considering their documented limitations, as described below.

The Support Vector Machines (SVM) family of algorithms is highly effective for both classification and regression problems. The core idea of the SVM algorithm is to identify the optimal linear separator that divides data points into two distinct classes. To overcome modeling problems that arise from very complex data, possible outliers, noise, or non-linear patterns, SVMs employ the so-called "kernel trick", a technique that maps data from its original data space to higher-dimensional feature spaces, where accurate classification is possible Cortes & Vapnik (1995). Key advantages of SVMs include robustness to extreme values, high prediction accuracy, and strong generalization abilities with limited hyperparameter tuning. However, SVMs can perform poorly when the number of features (in the feature space) greatly exceeds the number of observations used for training Mouchtaris et al. (2021).

We also use the CART (Classification and Regression Trees) algorithm, often referred to as decision trees (Breiman et al. (1984)). This model uses the Gini impurity index -a measure of the probability that a randomly selected observation is misclassified- to create optimal splits and build decision trees. Although decision trees are easy to train, understand, and they perform very well in-sample, nonetheless, they are prone to overfitting, resulting in very poor performance in new data. Overfitting is a common manifestation of the well-known trade-off between bias (in-sample error) and variance (difference between in-sample and out-of-sample performance).

To address the limitations of decision trees, Breiman (2001) introduced the Random Forest algorithm. Random forests combine the concept of decision trees with the bagging technique (Breiman (1996)). During training, the algorithm constructs multiple decision trees, each based on a bootstrapped sample of the original dataset and a random subset of variables. This approach reduces overfitting and improves generalization. Random forests also provide insight into the relevant forecasting importance of the independent variables through its built-in Variable Importance Measure (VIM). Potential drawbacks include the computational resources required to train multiple trees and the complexity of interpreting the ensemble model.

XGBoost (eXtreme Gradient Boosting, Chen & Guestrin (2016)) is another powerful ensemble learning algorithm. Unlike Random Forests, it builds decision trees sequentially, with each new tree designed to correct errors made by the previous trees through a process known as boosting. XGBoost minimizes a specific objective function, typically combining a loss function with a regularization term, which improves both the in-sample accuracy and the generalization ability out-of-sample. While XGBoost offers significant advantages, including ease of implementation and bias reduction, its disadvantages include complex hyperparameter tuning and high computational requirements.

3.2 The Dataset

Our dataset is compiled of 9 variables, namely, GDP, government bond yields, government debt, aggregate consumption, inflation, ECB interest rate, employment, aggregate investment and wages. For each of these variables, we use up to 12 lags to address temporal dependencies, resulting in a 108-dimensional input vector. The target variable is the cyclical component of the state of the government debt, derived from the Hamilton filter. If this is above trend (positive) at t+1 , it is labeled as 1, and if it is below trend (negative), it is labeled as 0. According to this setting, our goal is to forecast whether the government debt will be above or below its long-run trend in the next period. In this manner, we are able to forecast unexpected short-run deviations of public debt and align economic policy accordingly.

From the 10,000 observations of artificial data obtained from the DSGE model, 80% are used as the training set (in-sample) for model training, while the remaining 20% are used as the validation set (out-of-sample) to evaluate the models' ability to generalize on new and unknown data (figure 3). This 80-20 split is a widely adopted practice in machine learning and is effective for both classification and regression tasks. The real data, spanning the period from 2003Q1 to 2023Q2, are used as a second out-of-sample set.

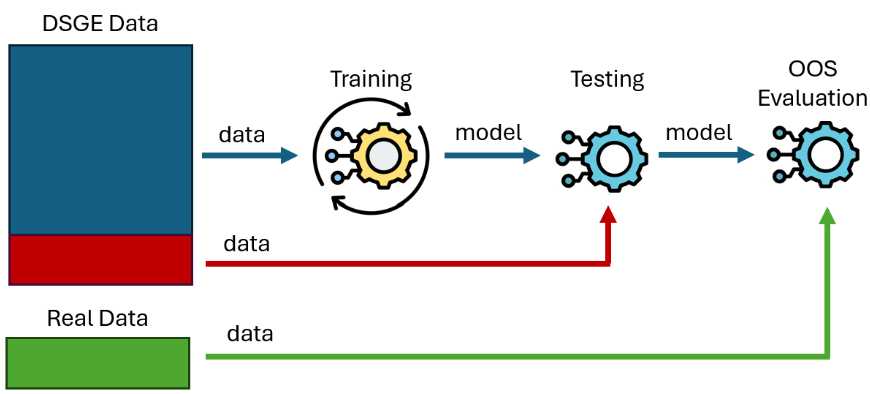

Figure 3: Use of data obtained from DSGE simulations

To further mitigate overfitting and ensure robust model selection, we use a 5-fold cross-validation technique in the training part of the DSGE dataset. By using cross-validation, the in-sample data are divided into 5 equally sized folds (subsets). For each configuration set of the hyperparameters, model training and testing are repeated 5 times. In each iteration, one fold is used for testing and the remaining four folds are used for training. The overall performance of each hyperparameter configuration is then calculated as the average performance across all five test folds.

This rigorous training and validation strategy helps to accurately optimize the hyperparameters while reducing the risk of overfitting, ensuring that the models perform efficiently when applied to new, unseen data.

To address the challenge of identifying the most relevant features for the prediction model, we applied the Sequential Feature Selection (SFS) technique, a robust technique designed to streamline the feature set for models that do not inherently perform feature selection. SFS is a greedy algorithm that starts with a single feature and sequentially adds the most important features, one at a time, to improve the performance of the model. The feature added at each iteration is selected based on the cross-validation score. Finally, the set of features with the highest cross-validation score is selected.

In our study, SFS was applied only to support vector machine (SVM) models, which have no built-in feature selection mechanisms. For tree-based models, explicit feature selection is not necessary, as these algorithms rank and order variables based on their importance during the training process, employing metrics such as the Gini significance.

3.3 Results

A 5-fold cross-validation approach was used to optimize the hyperparameters of each model. The SVM models were coupled with three kernels: the linear kernel, the Radial Basis Function (RBF) kernel and the polynomial kernel.

Table 4: Variables selected with SFS for every SVM kernel

#	**Linear kernel**	**RBF kernel**	**Polynomial kernel**
1	GDP_lag_1	Government yields_lag_1	Government yields_lag_1
2	Government yields_lag_1	Government debt_lag_1	Government debt_lag_1
3	Government debt_lag_1	Inflation_lag_1	Inflation_lag_1
4	Inflation_lag_1	GDP_lag_7	ECB_rate_lag_1
5	ECB_rate_lag_1	Government yields_lag_6	Investment_lag_1
6	GDP_lag_8	Government yields_lag_7	Wages_lag_1
7	GDP_lag_9	Consumption_lag_8	Government debt_lag_3
8	Government yields_lag_2	Inflation_lag_12	Consumption_lag_3
9	Government yields_lag_3	ECB_rate_lag_9	Inflation_lag_2
10	Government debt_lag_2	ECB_rate_lag_10	Inflation_lag_3
11	Consumption_lag_3		Inflation_lag_4
12	Inflation_lag_7		Inflation_lag_6
13	Inflation_lag_9		Inflation_lag_10
14	Inflation_lag_12		Inflation_lag_11
15	ECB_rate_lag_2		ECB_rate_lag_2
16	ECB_rate_lag_8		ECB_rate_lag_3
17	ECB_rate_lag_10		ECB_rate_lag_4
18	ECB_rate_lag_12		ECB_rate_lag_5
19	Employment_lag_2		ECB_rate_lag_7
20	Employment_lag_10		ECB_rate_lag_8
21	Investment_lag_4		Wages_lag_6
22	Investment_lag_9		Wages_lag_7
23	Wages_lag_4		Wages_lag_8
24	Wages_lag_5		Wages_lag_10
25	Wages_lag_6		

As a first step, we performed the SFS technique on the SVM models, separately, for every kernel tested, using accuracy to select the optimal number of variables. The selected variables are presented in (4).

Table 5: Results for In-sample, out-of-sample (OOS) and real data (real) results.

Metric	DT	RF	XGBoost	Linear	RBF	Poly
In-sample Acc	98.14%	98.47%	98.67%	99.06%	98.57%	97.13%
OOS Acc	98.05%	98.20%	98.25%	98.55%	98.00%	96.55%
OOS F1	97.09%	97.32%	97.39%	97.83%	97.03%	94.78%
Real Acc	88.57%	85.71%	90.00%	81.43%	85.71%	75.71%
Real F1	88.24%	84.85%	89.55%	80.00%	86.11%	75.36%

As mentioned in the methodology, tree-based methods rank and prioritize variables based on their importance during the training process; thus, they implicitly perform feature selection.

The metrics used for the evaluation of the models are the accuracy and the F1 score. The results are presented in Table (5).

Table 6: Confusion matrices for the real dataset for the best models from each algorithm.

Model	Predicted Class 0	Predicted Class 1
SVM Linear		
True Class 0	31	5
True Class 1	8	26
SVM RBF		
True Class 0	29	7
True Class 1	3	31
SVM Polynomial		
True Class 0	27	9
True Class 1	8	26
Decision Tree (DT)		
True Class 0	32	4
True Class 1	4	30
Random Forest (RF)		
True Class 0	32	4
True Class 1	6	28
XGBoost		
True Class 0	33	3
True Class 1	4	30

XGBoost emerges as the strongest performer in real-world data, achieving the highest accuracy (90.00%) and F1-score (89.55%). However, in out-of-sample (OOS) evaluations, SVM linear model exhibits marginally better performance, achieving the highest OOS accuracy (98.55%) and F1-score (97.83%). These results suggest that XGBoost excels in handling real-world complexity. This is also evident from the confusion matrices presented in Table (6) showing that the XGBoost model outperforms the competition in terms of accuracy for both classes.

In summary, XGBoost demonstrates a clear advantage in real-world classification settings.

4 Conclusion

This paper has demonstrated that combining DSGE models with machine learning techniques offers a promising approach to overcoming the challenges of public debt forecasting, especially in data-constrained contexts. By generating artificial data from an estimated DSGE model, we successfully trained four machine learning algorithms that achieved high accuracy and reliability. The results highlight the potential of this hybrid approach for enhancing macroeconomic forecasting. The results highlight the potential of this hybrid approach to improve economic forecasting, offering policymakers novel robust tools to evaluate public debt sustainability and forecast macroeconomic risks.

Importantly, this study represents a substantial step towards the sequential use of DSGE models and machine learning, setting the stage for future research to explore their complementary strengths. Future work could extend this framework to more complex macroeconomic environments, incorporate real-time data, or refine the machine learning techniques to improve both interpretability and robustness.

Funding: Emmanouil Sofianos would like to acknowledge that this work is part of the Interdisciplinary Thematic Institute MAKErS of the ITI 2021-2028 program of the University of Strasbourg, CNRS and INSERM. It has received financial support from the IdEx Unistra (ANR-10-IDEX-0002), and from the "Programme Investissement d'Avenir" as part of the SFRI-STRAT'US project(s) (ANR-20-SFRI-0012).

.

References

An, S. & Schorfheide, F. (2007), 'Bayesian Analysis of DSGE Models', Econometric Reviews **26**(2-4), 113–172.
URL: *https://ideas.repec.org/a/taf/emetrv/v26y2007i2-4p113-172.html*

Bachleitner, A. & Prammer, D. (2024), 'Don't blame the government!? an assessment of debt forecast errors with a view to the eu economic governance review', European Journal of Political Economy **82**.

Barbier-Gauchard, A. & Sofianos, M. (2024), 'Forecasting public debt in the euro area using machine learning: Decision tools for financial markets', BETA Working Paper (2024-47).

Belly, G., Boeckelmann, L., Mateo Caicedo Graciano, C., Di Iorio, A., Istrefi, K., Siakoulis, V. & Stalla-Bourdillon, A. (2023), 'Forecasting sovereign risk in the euro area via machine learning', Journal of Forecasting **42**(3), 453–738.

Breiman, L. (1996), 'Bagging predictors', Machine Learning **24**(2), 123–140.

Breiman, L. (2001), 'Random forests', Machine Learning **45**, 5–32.

Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. (1984), Classification and Regression Trees, Wadsworth. Inc. Monterey, California, USA.

Calvo, G. (1983), 'Staggered prices in a utility-maximizing framework', Journal of Monetary Economics **12**(1), 383–398.

Chen, T. & Guestrin, C. (2016), Xgboost: A scalable tree boosting system, in 'Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining', pp. 785–794.

Corsetti, G., Kuester, K., Meier, A. & Mueller, G. (2012), 'Sovereign risk, fiscal policy, and macroeconomic stability', IMF Working Paper (WP/12/33).

Cortes, C. & Vapnik, V. (1995), 'Support-vector networks', Machine Learning **20**(3), 273–297.

Di Bartolomeo, G., Di Petro, M., Saltari, E. & Semmler, W. (2018), 'Public debt stabilization: The relevance of policymakers' time horizons', Public Choice **177**(3-4), 287–299.

Estefania-Flores, J., Furceri, D., Kothari, S. & Ostry, J. (2023), 'Worse than you think: Public debt forecast errors in advanced and developing economies', Journal of Forecasting **42**(3), 685–714.

European Commission (2024), 'European economic forecast autumn 2024', Institutional paper 296 (november).

Gertler, M. & Karadi, P. (2011), 'A model of unconventional monetary policy', Journal of Monetary Economics **58**(1), 17–34.

Gogas, P. & Papadimitriou, T. (2021), 'Machine learning in economics and finance', Computational Economics **57**(1), 1–4.

Gogas, P., Papadimitriou, T. & Sofianos, M. (2021), 'Mind the gap: Forecasting euro-area output gaps with machine learning', Applied Economics Letters **29**(19), 1824–1828.

Gogas, P., Papadimitriou, T. & Sofianos, M. (2022), 'Forecasting unemployment in the euro area with machine learning', Journal of Forecasting **41**(3), 551–566.

Hamilton, J. D. (2017), Why You Should Never Use the Hodrick-Prescott Filter, NBER Working Papers 23429, National Bureau of Economic Research, Inc.
URL: *https://ideas.repec.org/p/nbr/nberwo/23429.html*

Hinterlang, N. & Hollmayr, J. (2022), 'Classification of monetary and fiscal dominance regimes using machine learning techniques', Journal of Macroeconomics **74**.

IMF (2024), 'Fiscal monitor'.

Kolasa, M. & Rubaszek, M. (2014), 'Forecasting with dsge models with financial frictions', Dynare Working Papers Series (Working Paper no. 40).

Lekhal, F. R. . M. (2024), 'Public external debt sustainability assessment: towards a machine learning based approach', Cogent Economics Finance **12**(1).

Mouchtaris, D. et al. (2021), 'Forecasting natural gas spot prices with machine learning', Energies **14**(18), 5782.

Sermpinis, G., Stasinakis, C., Theofilatos, K. & Karathanasopoulos, A. (2014), 'Inflation and unemployment forecasting with genetic support vector regression', Journal of Forecasting **33**(6), 471–487.

Silva, T., Wilhelm, P. & Amancio, D. (2024), 'Machine learning and economic forecasting: the role of international trade networks', Banco Central do Brasil Working Papers (597).

Smets, F. & Wouters, R. (2007), 'Shocks and frictions in us business cycles: A bayesian dsge approach', American Economic Review **97**(3), 586–606.

Stempel, D. & Zahner, J. (2023), 'Whose inflation rates matter most? a dsge model and machine learning approach to monetary policy in the euro area', Working paper series / Institute for Monetary and Financial Stability .