BANK OF GREECE
EUROSYSTEM

Working Paper

Stochastic debt sustainability analysis: a methodological note

Dimitrios Papaoikonomou

338

MARCH 2025

BANK OF GREECE
Economic Analysis and Research Department – Special Studies Division
21, E. Venizelos Avenue
GR-102 50 Athens
Tel: +30210-320 3610
Fax: +30210-320 2432

www.bankofgreece.gr

Published by the Bank of Greece, Athens, Greece

ISSN: 2654-1912 (online)
DOI: https://doi.org/10.52903/wp2025338

STOCHASTIC DEBT SUSTAINABILITY ANALYSIS: A METHODOLOGICAL NOTE

Dimitrios Papaoikonomou

Bank of Greece

ABSTRACT

This paper mainly focuses on the approach taken at the Bank of Greece regarding the application of stochastic methods to debt sustainability analysis, providing also a discussion of alternative options. Caution is advised in the way that stochastic methods are made operational, as they are far from exact and rely on assumptions of various degrees of plausibility, which are often not stated explicitly. A Monte Carlo exercise reveals that under the approach taken by the European Commission, the measurement of dispersion can be subject to significant bias, ranging from an over-estimation by 45% to an under-estimation in excess of 80%, depending on the time-series properties of the data.

JEL classification: H68, C53.

Keywords: Debt sustainability, stochastic simulations.

Acknowledgements: I have benefited from bilateral discussions with D. Papageorgiou and D. Louzis, as well as with other participants at an internal presentation at the Bank of Greece Economic Analysis and Research Department on March 4th 2024. I am also very grateful to the participants and organizers of the WGPF workshop on SDSA on October 9th 2024 at the ECB. I am most grateful to D. Louzis for very constructive comments on an earlier draft. Any remaining errors are my own.

Disclaimer: The views presented are those of the author and do not necessarily reflect the views of the Bank of Greece.

Corresponding author:
Dimitrios Papaoikonomou
Economic Analysis and Research Department
Bank of Greece
21, E. Venizelos Avenue
Athens 102 50, Greece
Tel. +30 210 3203827
email: dpapaoikonomou@bankofgreece.gr

1. Introduction

Interest on stochastic debt sustainability analysis (SDSA) has increased in recent years with SDSA modules forming an integral part of most major international organizations' toolkits. This is particularly relevant in the context of the European Union, as the new EU fiscal framework which has been in force since the end of April 2024 explicitly provides for an elevated role for SDSA. In particular, under the new framework Member States' fiscal targets will need to ensure that at the end of the medium term budgetary plans the debt to GDP ratio remains on a "plausibly downward path", as assessed by the European Commission's (EC) stochastic DSA approach.[1] This places stochastic DSA at the heart of the policy-making process, which in turn, becomes subject to the relative merits and limitations of the methods employed.

Debt sustainability analysis has traditionally relied on deterministic projections of debt to GDP in order to illustrate the sensitivity of the baseline scenario to changes in underlying assumptions about the evolution of debt drivers, such as fiscal outcomes, economic performance and borrowing cost. Stochastic DSA aims instead at providing a complete description of permissible outcomes, on the basis of historical experience. In the special case when the baseline arises as a forecast from an estimated econometric model, the task is reduced to a standard depiction of uncertainty around an econometric forecast. However, typically, in DSA analyses the baseline trajectory is itself deterministic in the sense that it is given exogenously, drawing on various different sources of information. Hence, the question arises of how to assign stochastic properties to a deterministic baseline?

At the risk of over-simplifying, available options can be grouped into two broad categories. Table 1 provides an overview, summarizing the main features, including caveats and an indicative list of institutions that is by no means exhaustive. The first broad category is labelled "Model-based" and relies on the estimation of an empirical model of debt drivers. The estimated model is used in order to generate a simulated distribution of debt trajectories, the properties of which are applied around the deterministic DSA baseline. The model-based approach is followed by the UK Office for Budget Responsibility (OBR), who provide very clear documentation in OBR (2021) and Steel (2021). Their method constitutes the closest analogue to the approach

[1] Articles 6, point (a) and 10(1) of Regulation (EU) 2024/1263 and European Commission (2024), Annex II.1.

employed at the Bank of Greece. The UK OBR present also an interesting example in that they have recently switched over to the model-based approach as their primary method. A similar approach is also employed by the ECB, albeit with the notable difference that the simulated distribution is not applied to an exogenously given deterministic baseline.[2] The model-based approach is grounded on the residual-based bootstrap which has a long tradition in econometric practice.[3] The validity of the method hinges on a sufficiently well-specified econometric approximation of the underlying data, typically a VAR. In multi-country applications it may not be straightforward to evaluate the extent to which differences in the simulated distributions across countries reflect genuine differences in uncertainty, or are driven by model specification. This can be an undesirable feature when the primary focus is on cross-country comparison. Furthermore, the alignment of the simulated distribution with the deterministic baseline may rest on arbitrary assumptions.

The second broad category is labelled "Raw data" and it uses available data points directly in order to obtain measures of uncertainty. The properties of these measures are next used in order to simulate a large number of representative shocks, which are then applied to the baseline in order to derive a distribution of debt trajectories. This is the general approach taken by the EC and the IMF, with more extensive documentation available for the former.[4] In general, raw data can refer either to observations of past forecast errors, or to historical data on the actual series of macroeconomic and fiscal debt drivers. One practical limitation in using forecast errors is that they are typically available at relatively short horizons, whereas debt sustainability analyses usually stretch over several years. This limitation is particularly binding in the case of Greece, as the currently uniquely favourable debt structure demands that the analysis extend over several decades.[5] Of the major international institutions that apply the raw data approach, the EC and the IMF have both opted for using actual data on debt drivers. This escapes the need for country-specific models and the method can be applied

---

[2] See Bouabdallah et al. (2017).
[3] For an early literature review see Li and Maddala (1996).
[4] See Annex A4, pp. 139-145 in European Commission (2024).
[5] As a legacy of the EU/IMF programmes the bulk of Greek public debt consists of official sector loans provided under concessional terms that include grace periods, long maturities and interest deferrals. Exposure to risk can be expected to increase as official loans get gradually re-financed on market terms. Due to the long maturities involved, this process will take several decades to complete, the final repayment being due in 2070.

uniformly across a large number of countries.[6] One important caveat, however, is that macroeconomic time series are typically not white noise in that they tend to display high autocorrelation. This can complicate the simulation of representative samples and as is demonstrated in section 4b, it can introduce non-negligible bias.

Table 1: Different approaches to stochastic DSA

	Model based	Raw data
Methodology	a. Estimate an empirical model of debt drivers. b. Use model to simulate the distribution of debt trajectories. c. Apply the properties of the simulated distribution around the DSA baseline.	a. Obtain measures of uncertainty based on raw data. b. Use the properties of (a) to simulate a large number of shocks. c. Apply the simulated shocks to the baseline to derive a distribution of debt.
Institutions	• UK Office for Budgetary Responsibility (OBR) • ECB (?) • Bank of Greece	• European Commission • IMF • UK OBR (formerly)
Caveats	• Not well-suited for cross-country comparisons • Alignment with baseline rests on possibly arbitrary assumptions	• Limited availability of forecast errors over long horizons. • Raw data can be non-random, which may undermine/complicate the simulations in step (b).

2. Specification and estimation of the empirical model

2a. Debt accounting

As a starting point consider the following fundamental debt accounting identity:

$$\Delta D_t \equiv -BB_t + DDA_t \qquad (1)$$

where D_t is the stock of gross government debt in year t, $\Delta D_t = D_t - D_{t-1}$, BB_t is the budget balance, which in turn is equal to the primary balance PB_t minus interest payments INT_t, DDA_t stands for deficit-debt adjustment, also frequently called 'stock-flow adjustment' and all variables are measured in nominal terms (EUR millions). For simplicity we abstract from debt denominated in foreign currency, which can be considered negligible in the case of Greece and other euro area member states. The debt accounting identity simply states that changes in the stock of debt are either recorded through the budget balance (above the line) or through DDA_t (below the line).

Expressing this as a share of nominal GDP results in the familiar law of motion of the debt to GDP ratio:

$$\Delta d_t = -pb_t + dda_t + \left(\frac{iir_t - g_t}{1+g_t}\right) d_{t-1} \qquad (2)$$

where $iir_t = INT_t / D_{t-1}$ is the implicit interest rate, g_t denotes the nominal GDP growth rate and remaining lower case letters denote division by nominal GDP.

[6] Working with forecast errors, instead, typically requires country-specific econometric estimates of the parameters of the underlying distribution (usually a 2-piece Normal).

This expression indicates that it suffices to know pb_t, g_t, iir_t and dda_t in order to trace the evolution of d_t, which motivates the choice of the endogenous variables in the specification of the empirical model below.

2b. A simple BVAR model of debt drivers

Assuming that the debt drivers are stochastic processes with finite mean and variance, they are approximated by a finite order VAR of the following form:

$$\mathbf{y}_t = A_1\mathbf{y}_{t-1} + B_0\mathbf{x}_t + C_0\mathbf{z}_t + \mathbf{e}_t \tag{3}$$

where $\mathbf{y}_t = [pb_t, g_t, iir_t, dda_t]'$ is an endogenous vector containing the four debt drivers that define the law of motion of the debt to GDP ratio, namely: the primary balance as a share of GDP (pb_t), the nominal GDP growth rate (g_t), the implicit interest rate (iir_t) and deficit-debt adjustment as a share of GDP (dda_t). Vector $\mathbf{x}_t$ contains exogenous variables and deterministic terms controlling for country-specific features related to the EU/IMF financial assistance to Greece during the sovereign debt crisis. In particular, $\mathbf{x}_t$ = [psi_t, $bankingsupport_t$, $kcontrols_t$, $buffer_t$, ($pre\_MoU_t * pb_{t-1}$)]', where psi_t is a binary dummy variable controlling for the Private Sector Involvement in the sovereign debt restructuring in 2012, $bankingsupport_t$ denotes the GDP share of public expenditure in support of the financial sector, $kcontrols_t$ accounts for the imposition of capital controls in 2015, $buffer_t$ controls for the ESM loan for building the cash buffer in 2018 and pre_MoU_t controls for the years prior to 2010 and interacts with lagged primary balance in order to permit a regime shift in the dynamics of fiscal performance following the introduction of the EU/IMF programmes. Finally, $\mathbf{z}_t = [1, year2020_t, year2021_t]'$ includes the intercept and controls for the effects of the COVID-19 pandemic through self-explanatory 0/1 dummies. The vector $\mathbf{e}_t$ collects the reduced-form residuals with $E(\mathbf{e}_t) = 0$ and $E(\mathbf{e}_t\mathbf{e}_t{}') = \mathbf{\Omega}$ positive definite.

A number of observations are in order. First, dda_t is frequently treated as a non-stochastic process set to zero. The UK OBR do generate forecasts for dda_t, but do not include it among the endogenous variables in their VAR model. The European Commission do not allow for uncertainty associated with dda_t and in general, the same applies also to the IMF.[7] Our decision to explicitly include dda_t among the endogenous

[7] The IMF may exceptionally allow for uncertainty related to dda_t. See footnote 49, p. 46 in IMF (2022).

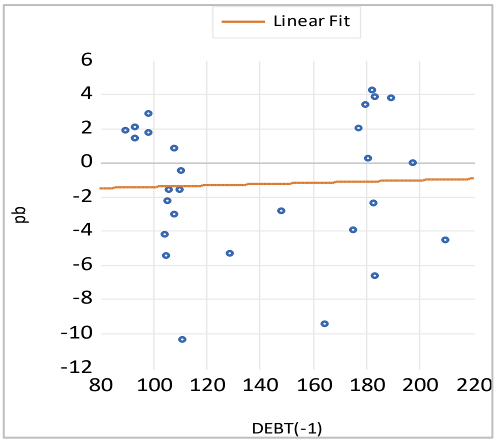
variables is guided by the fact that in the case of Greece dda_t has displayed sizeable variation and can be a considerable source of uncertainty, as pointed out also by the Hellenic Fiscal Council (2022).[8] A more sinister argument in favour of treating dda_t as an endogenous variable is that this would account also for the possibility that dda_t is affected by incentives for window-dressing fiscal outcomes.[9]

Second, the model does not include lagged debt, as suggested in the influential work by Favero and Giavazzi (2007). One reason is that in the case of Greece, there has been very little historical evidence of fiscal policy reacting to the level of debt. Chart 1 provides a telling scatter plot of the primary balance (pb_t) against lagged debt (d_{t-1}) for all available annual data points during 1996-2023. One further reason is that the inclusion of a debt feedback effectively imposes stationarity on the debt to GDP ratio. This is important in the context of policy evaluation, as it ensures well-behaved impulse responses. In the context of debt sustainability analysis, however, it makes little sense to *a priori* impose the very property one aims to assess. The UK OBR also do not include a debt feedback, although they do not explicitly discuss their decision.

Lastly, it should be stressed that the model is not intended as a forecasting tool. At this stage it is assumed that there already exists a deterministic baseline representing one's best guess of the evolution of the debt drivers. Instead, the purpose of the model is to shed light on the properties of the random shocks that have affected the determinants of public debt on the basis of past experience. In other words, the model essentially functions as a filter for removing the systematic, non-random element of the debt drivers $\mathbf{y}_t$ in order to reveal the residual vector $\mathbf{e}_t$.

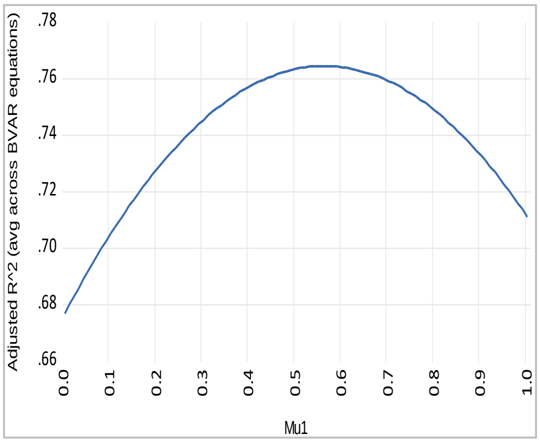
[8] Special Feature V, p.114 (only available in Greek).
[9] See von Hagen and Wolff (2006).

Chart 1: Historical scatter plot of pb_t against d_{t-1} (1996-2023)



Notes: Correlation = 0.04

Chart 2: Adjusted R^2 as a function of the autoregressive hyper-parameter Mu1



Notes: The vertical axis measures the simple average of the Adjusted R^2 in the four BVAR equations

The model is estimated over the whole set of available annual data points using Bayesian methods, which are less susceptible to the VAR 'curse of dimensionality'. Data definitions and sources are summarized in Appendix 1. All estimation is carried out in Eviews12. We use Minnesota priors with all hyper-parameters at their default values, except for the AR(1) hyper-parameter Mu1, which is guided by the following considerations. Our priors, supported by unit root tests, clearly point to stationarity and we expect stronger mean-reversion in dda_t, intermediate persistence in pb_t and g_t and significant persistence in the slow-moving iir_t. A choice for Mu1 in the region of 0.5 seems appropriate for reflecting stationarity with persistence and appears also broadly consistent with an optimal average fit, as illustrated in Chart 2. The value of Mu1 was further fine-tuned to 0.47 to better align the long-term properties of the BVAR with baseline assumptions.[10] Details on the estimated coefficients and the reduced-form residuals are provided in Table 2.

[10] The long-term properties of the BVAR can be sensitive to choices in priors and model specification. While the alignment of the BVAR properties with baseline assumptions is a welcome feature, it is not a pre-requisite for the validity of the stochastic analysis. The sensitivity of the dispersion of the simulated distribution of debt to different choices of Mu1 is illustrated in Appendix 3.

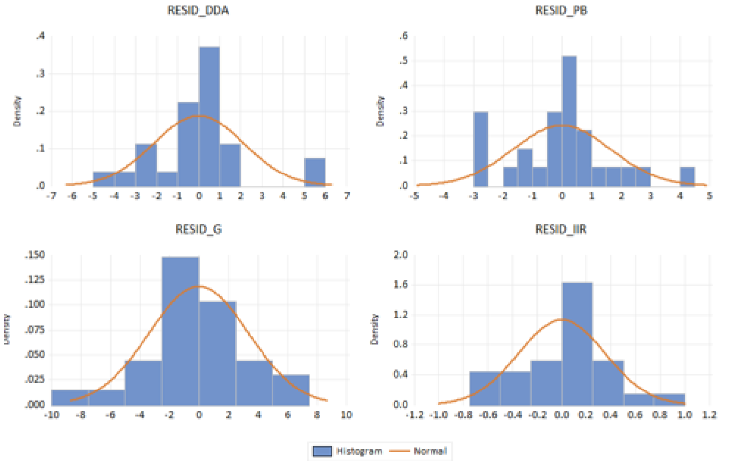
Table 2: BVAR estimates

A. Coefficients

Sample (adjusted): 1997 2023
Prior type: Litterman / Minnesota
Hyper-parameters: Mu1: 0.47, L1: 0.1, L2: 0.99, L3: 1, L4: inf
Standard errors in ( )

	DDA	PB	G	IIR
DDA(-1)	0.320396	-0.050206	0.028174	-0.004189
	(0.08259)	(0.08674)	(0.13885)	(0.01439)
PB(-1)	-0.019352	0.463962	0.086103	-0.001221
	(0.08242)	(0.08782)	(0.13952)	(0.01446)
G(-1)	-0.019660	-0.000577	0.535321	0.002467
	(0.05239)	(0.05540)	(0.08941)	(0.00919)
IIR(-1)	0.169079	-0.450332	-0.065711	0.791827
	(0.20839)	(0.22035)	(0.35281)	(0.03661)
C	-0.177844	2.702228	2.112465	0.714382
	(1.23378)	(1.30486)	(2.08880)	(0.21669)
PSI	-40.07687	-0.205966	-5.528772	-1.415818
	(2.86333)	(3.02553)	(4.84539)	(0.50204)
BANKINGSUPPORT	-0.596767	0.987460	-0.030869	0.062066
	(0.37151)	(0.39136)	(0.62661)	(0.06494)
KCONTROLS	-10.40221	-1.619167	-1.762008	-0.345703
	(2.72751)	(2.88366)	(4.61680)	(0.47852)
BUFFER	10.05417	0.950885	-1.489799	-0.184368
	(2.73667)	(2.89423)	(4.63283)	(0.48021)
PRE_MOU*PB(-1)	0.134095	1.079138	0.392924	-0.011456
	(0.30314)	(0.32060)	(0.51319)	(0.05319)
@YEAR=2020	-2.655028	-10.08953	-13.21116	-0.516920
	(2.73652)	(2.89423)	(4.63265)	(0.48020)
@YEAR=2021	0.512624	-3.175886	13.88126	-0.549087
	(2.78574)	(2.94582)	(4.71764)	(0.48882)
R-squared	0.930634	0.828786	0.713490	0.977829
Adj. R-squared	0.879765	0.703228	0.503382	0.961571
Sum sq. resids	118.2208	70.81987	293.9342	3.220928
S.E. equation	2.807381	2.172861	4.426693	0.463388
F-statistic	18.29484	6.600858	3.395828	60.14268
Mean dependent	-0.922831	-1.364215	2.800689	4.097174
S.D. dependent	8.096287	3.988602	6.281569	2.363820

B. Residuals



Descriptives

	RESID_DDA	RESID_PB	RESID_G	RESID_IIR
Mean	-6.58E-17	1.30E-15	4.81E-16	2.52E-15
Median	1.78E-15	8.88E-16	-1.78E-15	2.66E-15
Maximum	5.740903	4.489933	7.255440	0.918115
Minimum	-4.042848	-2.616462	-7.943062	-0.630253
Std. Dev.	2.132359	1.650406	3.362315	0.351969
Skewness	0.888980	0.468455	-0.126492	0.334585
Kurtosis	4.789730	3.667157	3.248495	3.256844
Jarque-Bera	7.159813	1.488263	0.141469	0.577976
Probability	0.027878	0.475147	0.931709	0.749021
Sum	-8.88E-16	3.61E-14	8.88E-15	6.71E-14
Sum Sq. Dev.	118.2208	70.81987	293.9342	3.220928
Observations	27	27	27	27

Covariance matrix

	DDA	PB	G	IIR
DDA	4.378549	0.714879	-2.117441	0.042264
PB	0.714879	2.622958	0.097011	-0.205500
G	-2.117441	0.097011	10.88645	-0.121056
IIR	0.042264	-0.205500	-0.121056	0.119294

3. Stochastic simulation of debt and Gross Financing Needs trajectories

3a. Bootstrap design

Stochastic simulations of debt and GFN trajectories are based on 100k bootstrap draws from the full-sample residuals $\hat{\mathbf{e}}_t$ recovered from equation (3). This avoids the need for arbitrary assumptions regarding the properties of the shocks. Despite accounting for a variety of country-specific influences, a number of sizeable outliers remain. As a result, the properties of the bootstrap draws can vary depending on the sample period. Generally, we find that restricting bootstrap draws to the more recent observations after the great financial crisis results in smaller dispersion compared to drawing from the full-sample. As such, our choice to draw from the full sample should be understood as a conservative choice that introduces higher uncertainty compared to the more recent track record. Having said that, the intercept dummies controlling for the COVID-19 pandemic in 2020 and in 2021 effectively exclude the estimated effect from the set of permissible future shocks.

The bootstrap simulation of debt and GFN trajectories consists of the following steps:

Step 1: Estimate the BVAR model in equation (3) using the actual data in order to obtain the coefficient matrices $\hat{A}_1$, $\hat{B}_0$, $\hat{C}_0$ and the residual vector $\hat{\mathbf{e}}_t$.

Step 2: Draw randomly with replacement from $\hat{\mathbf{e}}_t$ in order to generate a large number of simulated shocks $\mathbf{e}_t^{(i)}$ for $t = 1, \dots, T$ and $i = 1, \dots, 100k$.

Step 3: Generate a large number of random draws $A_1^{(i)}$, $B_0^{(i)}$ and $C_0^{(i)}$, $i = 1, \dots, 100k$ from the posterior distribution of the estimated BVAR coefficients in step 1.

Step 4: Combine the simulated shocks in step 2 with the random draws from the BVAR coefficients in step 3 to generate a large number of simulated paths for debt drivers as: $\mathbf{y}_t^{(i)} = A_1^{(i)}\mathbf{y}_{t-1}^{(i)} + B_0^{(i)}\mathbf{x}_t^{(i)} + C_0^{(i)}\mathbf{z}_t^{(i)} + \mathbf{e}_t^{(i)}$ for $t = 1, \dots, T$ and $i = 1, \dots, 100k$.

Step 5: For each of the simulated $\mathbf{y}_t^{(i)}$s in step 4 generate debt trajectories as:

$$\Delta d_t^{(i)} = -pb_t^{(i)} + dda_t^{(i)} + d_{t-1}^{(i)}(iir_t^{(i)} - g_t^{(i)})/(1 + g_t^{(i)}) \quad (4)$$

GFN trajectories are similarly simulated under baseline assumptions for the maturity profile of new issuance[11] and conditionally on information on the maturity of the historically accumulated debt stock. In particular:

$$gfn_t^{(i)} = am_t^{(i)} + \Delta d_t^{(i)} \quad (5)$$

where $gfn_t^{(i)}$ is the GFN to GDP ratio generated using the *i*-th bootstrap draw and $am_t^{(i)}$ is the GDP share of amortization payments, including the stock of short-term debt that is refinanced each year.

3b. Simulated uncertainty around the baseline

Following steps 1-5 described above generates empirical distributions of 100k simulated debt and GFN trajectories that are representative of the historical track record, as described by the simple BVAR model. Chart 3, panel A plots the resulting fan charts along with the simulated mean (grey), median (green), a Kernel approximation of the mode (blue) and the deterministic baseline (black line).[12] One observes that the simulated distributions are not symmetric, as the mean exceeds the

[11] The baseline assumption is that new issuance has an average maturity of approximately 6,5 years, which is broadly in line with the historical record before the Great Financial Crisis.
[12] The simulated trajectories of individual debt drivers can be found in Appendix 2.

median. This appears to be less important at short horizons, but becomes more pronounced over time. The reported histograms plot the simulated distributions for the year 2060, which are clearly positively skewed, reflecting sizeable positive outliers.

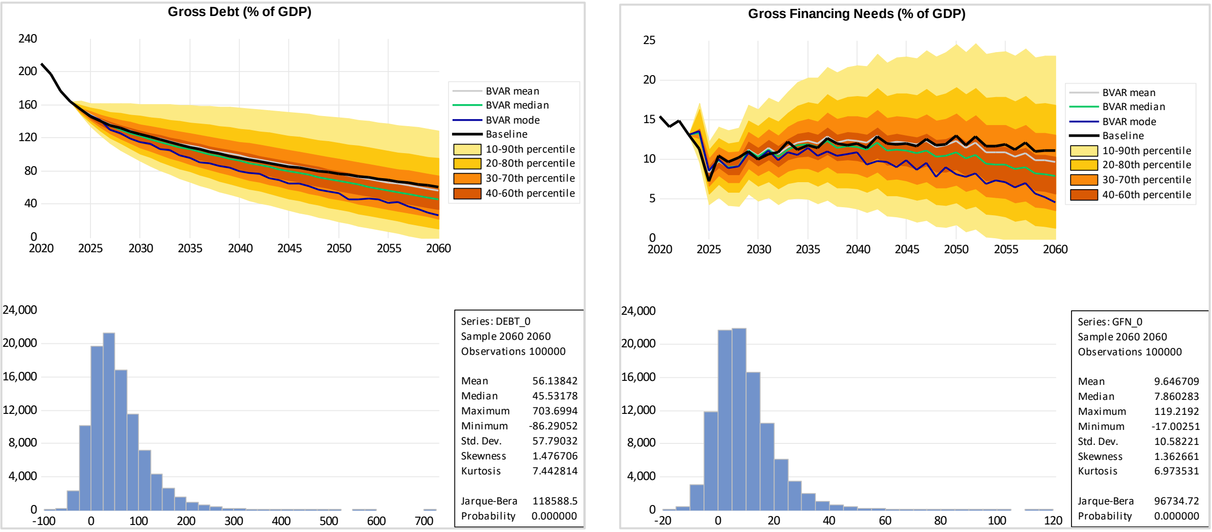
To obtain a distribution of risks around the baseline we align the simulated mean trajectory with the DSA baseline by mechanically shifting the simulated trajectories by the difference between the baseline and the simulated mean, so that $aligned_t^{(i)} = notaligned_t^{(i)} + baseline_t - mean_t$ for all $i = 1, \dots, 100k$. The decision to treat the baseline as a mean forecast differs from the OBR, who treat the baseline as a median forecast, instead.[13] For a symmetric distribution of future outcomes the median would coincide with the mean, yielding an unbiased forecast with zero forecast errors on average. As was already established, however, the simulated distributions are not symmetric, but positively skewed. Given this asymmetry, treating the baseline as a median forecast would be equivalent to assuming that the baseline is systematically biased, in which case it would make more sense to revise the baseline. If one considers instead the baseline not to be systematically biased, the asymmetry of the simulated distributions requires that it represents the mean of future outcomes. In deciding to interpret the baseline as a mean forecast we draw confidence from the fact that it is already closely aligned with the simulated mean. This is however not a general result and we return to this in section 3c.

In the context of a positively skewed distribution a mean forecast implies more frequent over-estimation of moderate size, in order to compensate for rare but sizeable tail events. This is reflected in the asymmetric placement of the baseline trajectories covering the greater mass of the distributions in Chart 3, panel B. The simulated distributions can be used in order to quantify a number of probability metrics. Table 3 reports the evolution through time of the probability that specific adverse events materialize under baseline assumptions.

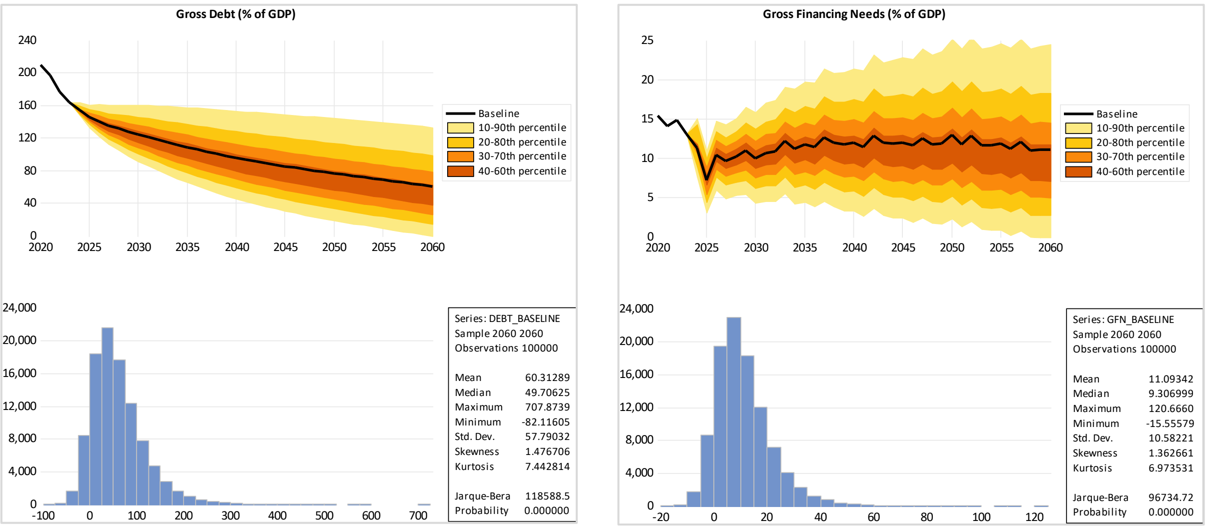
[13] The OBR differ also in employing a more sophisticated alignment method. See para. 3.20 in Steel (2021).

Chart 3: Simulated trajectories of debt/GDP (left) and GFN/GDP (right)

A. Before aligning the simulated mean with the baseline



B. After aligning the simulated mean with the baseline



Notes: Based on 100k BVAR simulations. The histograms plot the simulated distributions in year 2060.

Table 3: Probability metrics after alignment with baseline

	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2050	2060
1. Pr(debt rule violation)	11.0	22.9	36.1	36.8	41.3	41.4	42.2	42.7	42.7	42.7	42.9	42.9	42.7	42.7	42.7	42.7	42.7	41.7	41.0
2. Pr($d_{t+h} > d_t$)	11.0	6.5	7.5	7.5	8.0	8.2	8.4	8.5	8.5	8.3	8.2	8.1	7.9	7.7	7.6	7.4	7.3	6.4	5.4
3. Pr($d_{t+h} > 90$)	100.0	100.0	100.0	99.9	98.8	95.7	91.2	86.1	80.8	75.5	70.7	66.3	62.1	58.5	55.1	52.1	49.2	32.9	24.0
4. Pr($d_{t+h} > 60$)	100.0	100.0	100.0	100.0	100.0	100.0	99.9	99.6	98.8	97.5	95.6	93.3	90.6	87.8	84.8	81.9	79.1	56.7	42.0
5. Pr($GFN_{t+h} > 15$)	11.0	0.1	9.1	6.7	10.2	17.1	13.4	18.1	19.4	27.7	23.8	27.3	25.6	32.3	29.2	28.7	29.7	35.0	28.7
6. Pr($GFN_{t+h} > 20$)	0.0	0.0	0.1	0.2	0.6	2.4	1.9	3.7	4.3	7.8	7.4	9.4	9.2	13.2	11.7	12.0	13.0	19.3	16.6

Notes: Row 1 reports the probability that debt/GDP does not decline year-on-year by at least 1pp when debt is above 90% of GDP and by at least 0.5pp when debt is between 60% and 90% of GDP. Rows 2-4 report the probability that debt/GDP exceeds (a) the latest historical value, (b) 90% and (c) 60%, respectively. Rows 5 & 6 report the probability that GFN/GDP exceeds 15% and 20%, respectively. The following colour code applies: <10%, 10-20%, 20-30%, 30-40%, 40-60%, 60-70%, 70-80%, 80-90%, > 90%.

3c. The baseline as a long-run steady state

As argued above, the decision to treat the baseline as a mean forecast reflects the skewness of the simulated distribution, which would render a median forecast biased. Yet, the interpretation of the baseline as an unbiased forecast is in general arbitrary. In the particular case considered here, the mean forecast generated from the estimated BVAR is well-aligned with our baseline. This is a welcome feature demonstrating that the baseline can arise as an unbiased forecast under a plausible set of model specifications. However, this is not true for every plausible set of modelling choices. Indeed, small changes in the Bayesian hyper-parameters can lead to non-negligible deviations of the mean forecast from the baseline.

Rather than aligning the baseline with a specific metric of the simulated distribution, such as the mean, or the median, one could interpret instead the baseline as a long-run steady state from which debt drivers are permitted to deviate due to the simulated random shocks. In particular, the simulated debt drivers $\mathbf{y}_t^{(i)}$ generated in step 4 of the bootstrap procedure described in section 3a would assume the following general form:

$$\mathbf{y}_t^{(i)} = \left(I - \Lambda^{(i)}\right)\mathbf{y}_t^{baseline} + \Lambda^{(i)}\mathbf{y}_{t-1}^{(i)} + \mathbf{e}_t^{(i)} \quad (6)$$

where vector $\mathbf{y}_t^{baseline}$ collects the exogenously given values of the debt drivers under the deterministic baseline, $\Lambda^{(i)}$ is an autoregressive parameter matrix which could be informed from the BVAR estimate of $\hat{A}_1$ in equation 3 and $\mathbf{e}_t^{(i)}$ is a bootstrap draw from the BVAR residuals. The suggested specification escapes the need for aligning the baseline with a specific metric of the simulated distribution, thereby reducing the scope for arbitrary choices.

4. Commentary on the EC and IMF methods

4a. Normative overview

Under the revised EU fiscal framework the EC's stochastic DSA plays an important role in the design of national fiscal policies, as it is the method by which it is assessed whether the debt to GDP ratio remains on a "plausibly downward path" at the end of Member States' medium term budgetary plans. The method, which is described in detail in Annex A4 in European Commission (2024), looks at the historical correlations of the first differences of a number of variables that are required in order to trace the debt to GDP ratio, such as the GDP share of the primary balance (pb_t) and the GDP nominal growth rate (g_t). Stochastic shocks are then generated as random draws from a multivariate normal distribution with variance-covariance matrix equal to the one observed historically for the differenced variables. These shocks are then applied to the baseline values of the debt-driving variables to generate a range of debt outcomes around the baseline.

Darvas et al. (2023) recently noted a number of caveats, including the lack of clarity regarding the treatment of outliers and the potentially low number of random draws.[14] The Hellenic Fiscal Council has pointed out that the EC approach does not allow for the stochastic treatment of dda, which could be an additional important source of uncertainty in the case of Greece.[15] Yet, a more fundamental caveat seems to have escaped notice. The EC approach rests on the assumption that the first differences of variables such as pb_t and g_t are random variables that are jointly normally distributed. This is equivalent to assuming that the levels of these variables are random walks, i.e. unit root processes denoted $I(1)$.[16] Such a property implies unbounded variance, which is very difficult to defend in the case of pb_t, g_t, or interest rate variables, all routinely treated as stationary $I(0)$ processes instead. The first difference of a stationary process – unlike that of a unit root process – is not white noise, and as such, it cannot be treated as a random draw from a normal distribution.

The IMF method escapes this problem by drawing blocks of consecutive observations with replacement directly from the available historical series.[17] This process is known as block bootstrap and is aimed at retaining (part of) the autocorrelation structure. Yet, given that the time series on debt drivers are widely considered to be stationary, the procedure is subject to the critique of Politis and Romano (1994) that a block bootstrap of a stationary process is not itself stationary. Hence, simulating debt drivers by using the block bootstrap for resampling the actual data series will not retain a key statistical property observed in the historical record.[18]

---

[14] See Annex A.4.1.2, p. 58 in Darvas et al. (2023).

[15] Hellenic Fiscal Council (2022), Autumn Report, Special Feature V, p.114 (only available in Greek).

[16] For the n-vector $\mathbf{x}_t$ that includes variables like pb_t and g_t the EC assumes that $\Delta\mathbf{x}_t = \boldsymbol{\varepsilon}_t$, with $\boldsymbol{\varepsilon}_t \sim N(0,\Sigma)$, which is equivalent to assuming that $\mathbf{x}_t$ is a random walk since $\mathbf{x}_t = \mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t$.

[17] See section VI. para. 59 and footnote 51 in IMF (2022).

[18] The 'stationary bootstrap' developed by Politis and Romano (1994) is the earliest example of a more complex procedure that can address this limitation of the block bootstrap.

Sampling directly from time-series data poses a number of challenges to which model-based methods offer a comparatively simple solution. The main appeal of the raw-data approach is that it escapes the need for country-specific models and permits a uniform application across countries. This is an attractive feature in the context of cross-country comparisons, which lie at the core of the work of the EC and the IMF. The price, however, can be non-negligible, as indicated in the following subsection.

4b. Quantification of the dispersion bias in the EC method

As mentioned above, the European Commission's definition of the random shocks as the first difference of the debt drivers is equivalent to assuming that the debt drivers are unit root processes, which can be very difficult to defend in the case of pb_t, g_t, or interest rate variables, routinely treated as stationary processes, instead. To illustrate why this can be a source of bias, consider the case of a single debt driver x_t given by the following AR(1) process:

$$x_t = (1-\rho)\bar{x} + \rho x_{t-1} + \varepsilon_t \tag{7}$$

where $0 \leq \rho \leq 1$, $\bar{x}$ is a constant steady-state and $\varepsilon_t \sim iid\, N(0, \sigma^2)$. In the extreme case when $\rho = 0$ the AR(1) process collapses to $x_t = \bar{x} + \varepsilon_t$. Applying the European Commission's definition of a shock as Δx_t results in $\Delta x_t = \varepsilon_t - \varepsilon_{t-1}$. In this case $Var(\Delta x_t) = 2\sigma^2$, which clearly leads to an over-estimation of the true variance σ^2 of the random shocks ε_t. This bias is eliminated for $\rho = 1$ in which case x_t becomes a random walk with $\Delta x_t = \varepsilon_t$ and $Var(\Delta x_t) = Var(\varepsilon_t) = \sigma^2$.

There is, however, one additional source of bias related to the way that the shocks are applied to the deterministic baseline. Consider the simulated driver x_t^{sim} which is constructed as follows:

$$x_t^{sim} = x_t^{baseline} + \varepsilon_t^{sim} \tag{8}$$

where $x_t^{baseline}$ is the exogenously given value of x_t under the deterministic baseline and ε_t^{sim} is a random draw from a normal distribution with variance equal to $Var(\Delta x_t)$.[19] The absence of a lagged term means that x_t^{sim} permits no persistence

[19] The European Commission uses quarterly observations of x_t in order to generate the annual shocks ε_t^{sim} entering the simulation of x_t^{sim}, which is itself carried out at annual frequency. For simplicity it is assumed here that actual and simulated observations of x_t are of the same frequency.

which introduces the following trade-off. In order for the simulated shocks ε_t^{sim} to have a variance equal to the true variance σ^2, the driver x_t needs to be a random walk with full persistence, i.e. $\rho = 1$. Yet, the more this is the case, the less representative becomes the simulated x_t^{sim}, which by construction exhibits no persistence and thus under-states the true size of uncertainty. Hence, for $\rho = 0$ the persistence of the shocks is simulated accurately but their variance is over-estimated, while for $\rho = 1$ the variance of the shocks is accurately measured but their persistence is under-estimated.

The following Monte Carlo exercise has been carried out in order to quantify the bias involved in the measurement of dispersion of the simulated distribution.

Step 1: Generate a large number of AR(1) processes of length T as $x_t^{(i)} = (1-\rho)\bar{x} + \rho x_{t-1}^{(i)} + \varepsilon_t^{(i)}$, where $i = 1, \dots, 100k$ and $\varepsilon_t^{(i)}$ is a random draw from a $N(0, \sigma^2)$. The starting value is set at $x_0^{(i)} = 1$ and the parameters $\bar{x} = 1$ and $\sigma^2 = 0.01$. The total number of observations is set at $T = 66$, of which the first $T_{hist} = 29$ observations are treated as historical values. This corresponds to the length of the available historical record of annual data points for Greece spanning 1995-2023 and allows for a forecast horizon covering 2024-2060.

Step 2: For $t > T_{hist}$ compute $dispersion\_x_t$ as the difference between the 90th and the 10th percentiles of the empirical distribution of $x_t^{(i)}$ in step 1. This is the ‘true’ dispersion in line with the data generating process.

Step 3: For $t \leq T_{hist}$ compute the historical shocks as $\Delta x_t^{(i)}$ and for $t > T_{hist}$ generate the simulated shocks $\varepsilon_t^{sim,(i)}$ as random draws from a normal distribution with variance equal to $Var(\Delta x_t^{(i)})$.

Step 4: For $t > T_{hist}$ generate the simulated drivers $x_t^{sim,(i)} = x_t^{baseline} + \varepsilon_t^{sim,(i)}$, where $\varepsilon_t^{sim,(i)}$ is provided in step 3 and $x_t^{baseline}$ is an exogenously given baseline, which in this case is set equal to the mean of the empirical distribution of $x_t^{(i)}$ in step 1.

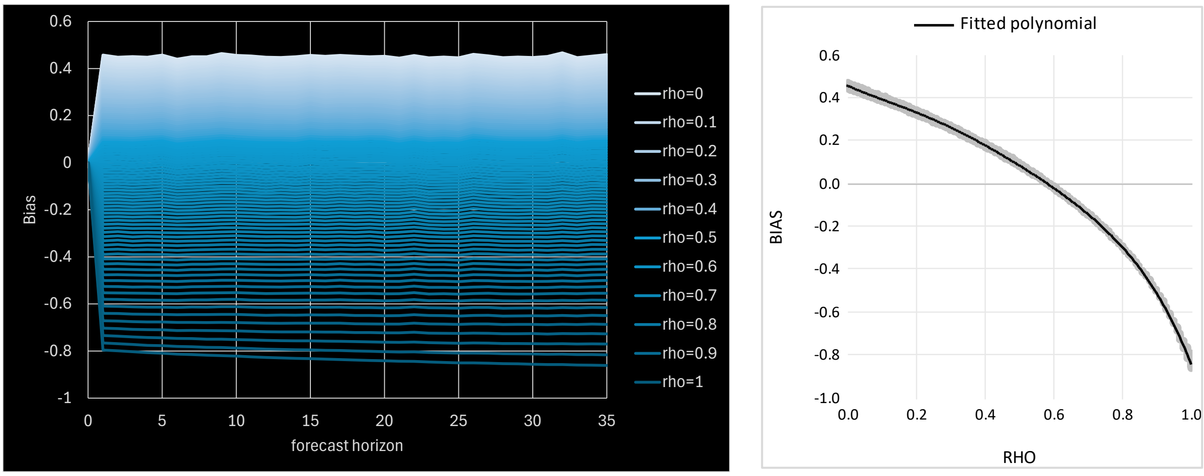
Step 5: Compute the simulated $dispersion\_xsim_t$ as the difference between the 90th and the 10th percentiles of the empirical distribution of $x_t^{sim,(i)}$ in step 4 and calculate $bias_t = \frac{dispersion\_xsim_t}{dispersion\_x_t} - 1$, using $dispersion\_x_t$ from step 2.

Step 6: Repeat steps 1-5 for different values of the autoregressive parameter ρ. In this case, steps 1-5 have been repeated 101 times starting from $\rho = 0$ and reaching $\rho = 1$ in increments of 0.01.

Chart 4 offers two alternative presentations of the results. Panel A plots the dispersion bias as a function of the forecast horizon, each horizontal line corresponding to a different value of the autoregressive parameter ρ. The bias in the measurement of dispersion remains broadly constant over different forecast horizons, but, as anticipated by the earlier discussion, it varies significantly with ρ. This is more clearly illustrated in panel B, which is a scatter plot of the dispersion bias over all horizons plotted against ρ. The bias is highest for $\rho = 0$ when dispersion is over-estimated by approximately 45%. As ρ increases the positive bias declines at an accelerated pace, turning negative for $\rho > 0.58$. When $\rho = 1$ dispersion is under-estimated by more that 80%. This is in line with the earlier discussion on the trade-off between over-estimation of the variance and under-estimation of persistence. Depending on their underlying statistical properties, debt drivers can be subject to very different bias.

Chart 4 – European Commission dispersion bias

A. As a function of the forecast horizon. **B. As a function of ρ.**



Notes: Dispersion is measured as the difference between the 90th and the 10th percentiles. Bias measures percentage deviation from 'true' dispersion. Rho is the autoregressive coefficient of the underlying AR(1) data generating process.

5. Concluding remarks

The primary aim of this paper is to present the approach taken at the Bank of Greece in applying stochastic methods to debt sustainability analysis, offering explicit discussion of the main conceptual and technical issues. Additionally, a discussion is provided of alternative methods employed by major institutions. Particular emphasis is placed on the European Commission method, motivated by its elevated role under the new EU fiscal framework and facilitated by the availability of detailed documentation. A Monte Carlo exercise reveals that the EC measurement of dispersion can be subject to significant bias, ranging from an over-estimation by 45% to an under-estimation in excess of 80%, depending on the time-series properties of the data. An important takeaway is that stochastic DSA is far from exact, which calls for caution in the way that such methods become operational for the purpose of policy assessment.

References

Bouabdallah, O., C. Checherita-Westphal, T. Warmedinger, R. de Stefani, F. Drudi, R. Setzet, A. Westphal (2017), "Debt sustainability analysis for euro area sovereigns: a methodological framework", ECB Occasional Paper Series, No. 185.

Darvas, Z., L. Welslau and J. Zettelmeyer (2023), "A quantitative evaluation of the European Commission's fiscal governance proposal", Working Paper 16/2023, Bruegel.

European Commission (2024), "Debt Sustainability Monitor 2023", European Economy Institutional Paper 271, March 2024.

Favero, C. and F. Giavazzi (2007), "Debt and the effects of fiscal policy", NBER Working Paper 12822.

Hellenic Fiscal Council (2022), Autumn Report.

International Monetary Fund (2021), "Review of the debt sustainability framework for market access countries", IMF Policy Paper No. 2021/003.

International Monetary Fund (2022), "Staff guidance note on the sovereign risk and debt sustainability framework for market access countries", Policy Paper No. 2022/039.

Li, H. and G. S. Maddala (1996), "Bootstrapping Time Series Models", *Econometric Reviews*, 15(2), 115-15.

Office for Budget Responsibility (2021), Economic and fiscal outlook, October.

Politis, D. N. and J. P. Romano (1994), "The stationary bootstrap", *Journal of the American Statistical Association*, Vol. 89, No. 428, pp. 1303-1313.

Steel, D. (2021), "Evaluating forecast uncertainty with stochastic simulations", Working paper No.17, Office for Budget Responsibility, December.

von Hagen, J. and G.B. Wolff (2006), "What do deficits tell us about debt? Empirical evidence on creative accounting with fiscal rules in the EU", *Journal of Banking & Finance* 30, pp. 3259–3279.

Appendices

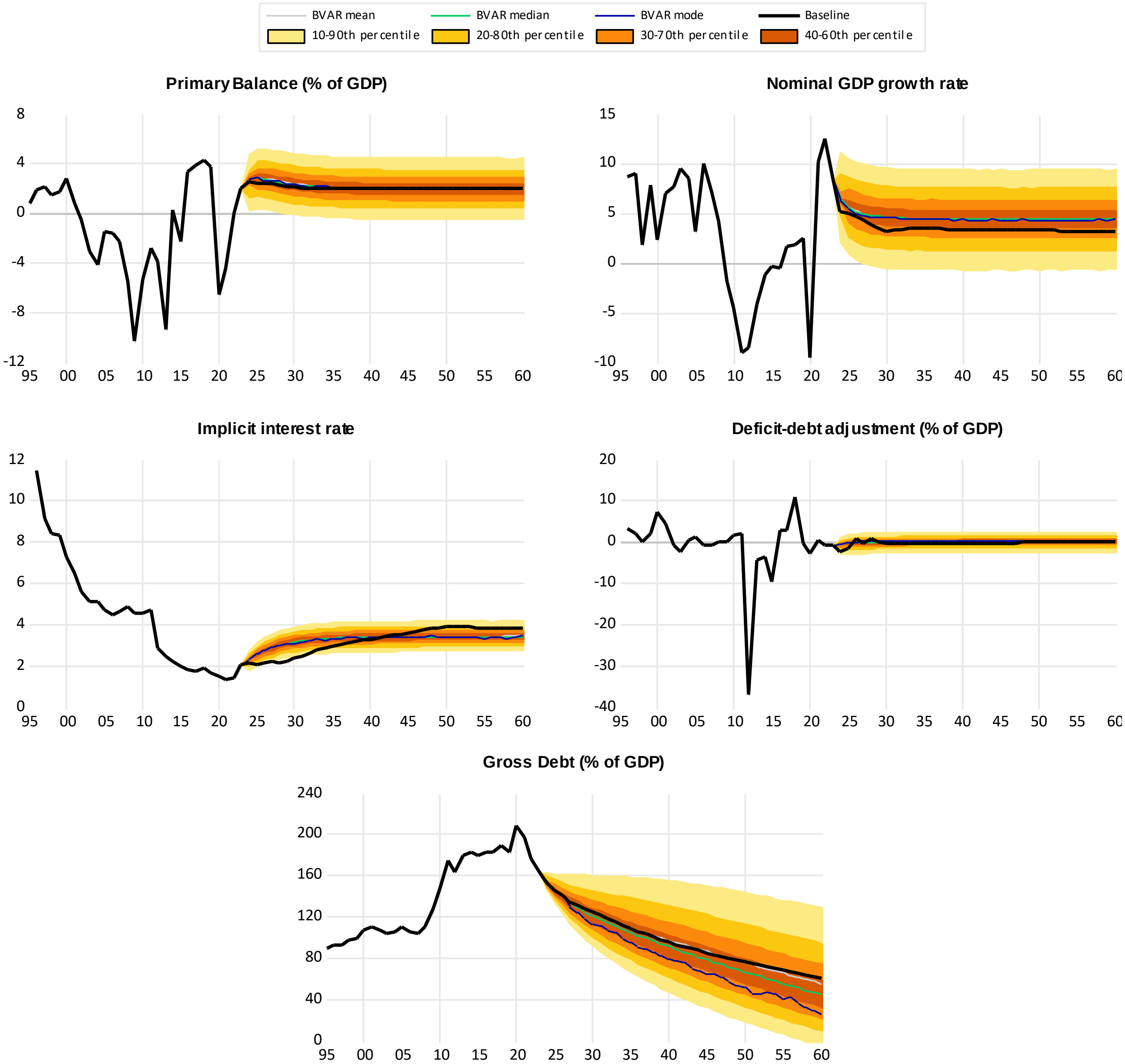
Appendix 1: Data sources and definitions

Frequency: Annual Sector: General Government Vintage: Autumn 2024 EDP	**Description**	**Source**	**Availability**
Primary data (in EUR million)			
D_t	Gross public debt	Eurostat	1995-2023
BB_t	Budget balance	Eurostat	1995-2023
INT_t	Interest expenditure	Eurostat	1995-2023
$PB_t = BB_t + INT_t$	Primary balance	Eurostat	1995-2023
$DDA_t = D_t - D_{t-1} + BB_t$	Deficit-debt adjustment	Eurostat	1996-2023
$BANKINGSUPPORT_t$	Banking support	Eurostat*	2007-2023**
GDP_t	Nominal GDP	Eurostat	1995-2023
AM_t	Amortization	PDMA	2024-2070
Variable transformations			
iir_t =100* INT_t/D_{t-1}			
$pb_t = 100 * PB_t/GDP_t$			
$dda_t = 100 * DDA_t/GDP_t$			
$g_t = 100 * (GDP_t - GDP_{t-1})/GDP_{t-1}$			
$bankingsupport_t = 100 * BANKINGSUPPORT_t/GDP_t$			
$am_t = 100 * AM_t/GDP_t$			

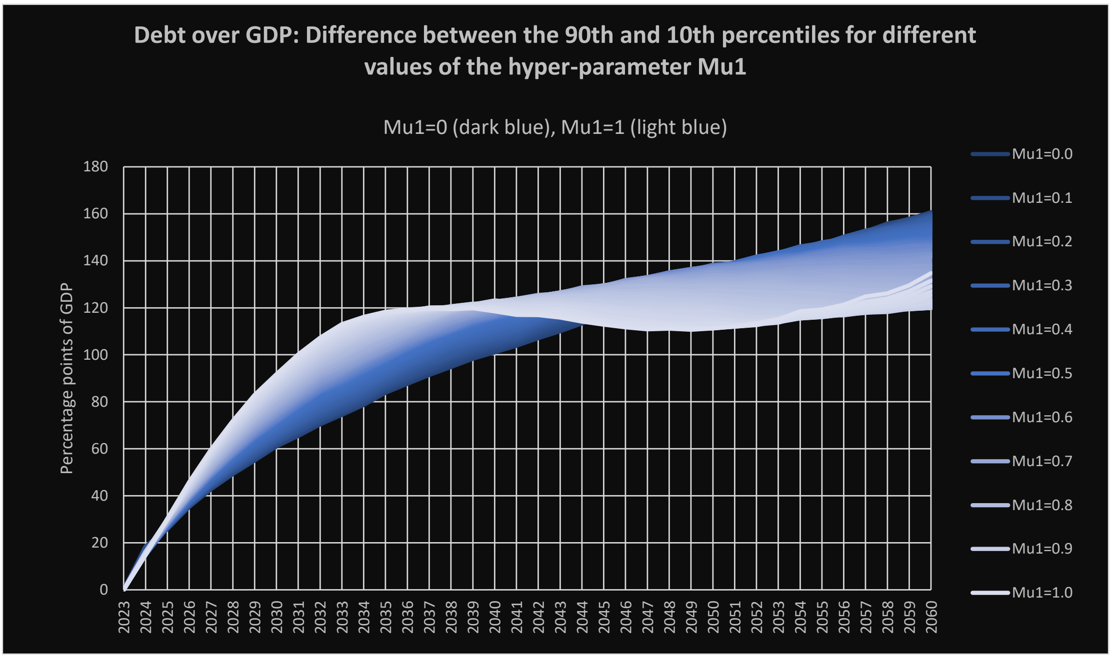
* Supplementary tables on government interventions to support financial institutions

** Set to zero before 2007.

Appendix 2: BVAR forecasts of debt drivers



Appendix 3: Sensitivity of dispersion to the autoregressive hyperparameter Mu1



BANK OF GREECE WORKING PAPERS

319. Dixon, H. T. Kosma, and P. Petroulas, "Explaining the endurance of price level differences in the euro area", May 2023.

320. Kollintzas, T. and V. Vassilatos, "Implications of market and political power interactions for growth and the business cycle II: politico-economic equilibrium", May 2023.

321. Bragoudakis, Z. and I. Krompas "Greek GDP forecasting using Bayesian multivariate models", June 2023.

322. Degiannakis, S. and E. Kafousaki "Forecasting VIX: The illusion of forecast evaluation criteria", June 2023.

323. Andreou C. P., S. Anyfantaki, C. Cabolis and K. Dellis, "Exploring country characteristics that encourage emissions reduction", July 2023.

324. Dimakopoulou, V., Economides, G., Philippopoulos, A., and V. Vassilatos, "Can central banks do the unpleasant job that governments should do?", December 2023.

325. Chrysanthakopoulos, C. and A. Tagkalakis, "The medium-term effects of fiscal policy rules", January 2024.

326. Manou, K. and E. Papapetrou, "Does uncertainty matter for household consumption? A mean and a two tails approach", February 2024.

327. Kakridis, A., "War, mobilization, and fiscal capacity: testing the bellicist theory in Greece, 1833-1939", March 2024.

328. Mavrogiannis, C. and A. Tagkalakis, "From policy to capital: assessing the impact of structural reforms on gross capital inflows", April 2024

329. Delis, P., S. Degiannakis, G. Filis, T. Palaskas and C. Stoforos, "Determinants of regional business cycle synchronization in Greece", May 2024.

330. Sideris, D. and G. Pavlou, "Market power and profit margins in the Euro area countries in the post-pandemic period", June 2024.

331. Kasimati, E. and N. Veraros, "The dry-bulk shipping market: a small econometric model", September 2024.

332. Mermelas, G. and A. Tagkalakis, "Monetary policy transmission: the role of banking sector characteristics in the euro area", November 2024.

333. Anastasiou, D., Pasiouras, F., Rizos, A., and A. Stratopoulou, "Do macroprudential policies make SMEs more-or-less discouraged to apply for a bank loan?", December 2024.

334. Malliaropulos, D., Passari, E., and F. Petroulakis, "Unpacking commodity price fluctuations: reading the news to understand inflation", December 2024

335. Degiannakis, S. and E. Kafousaki, "Disaggregating VIX", January 2025

336. Degiannakis, S., Delis, P., Filis, G., and G. Giannopoulos, "Trading VIX on volatility forecasts: another volatility puzzle?", February 2025

337. Papadopoulos, G., Ojea-Ferreiro, J., and R. Panzica, "Climate stress test of the global supply chain network: the case of river floods", February 2025