Center for Quantitative Economics

Stochastic debt sustainability analysis using time-varying fiscal reaction functions
An agnostic approach to fiscal forecasting

Tore Dubbert † #

104/2022

† Department of Economics, University of Münster, Germany
# Department of Business and Economics, University of Göttingen, Germany

Stochastic debt sustainability analysis using time-varying fiscal reaction functions

An agnostic approach to fiscal forecasting

Tore Dubbert [1] [2]

[1]University of Göttingen

[2]University of Münster

September 30, 2022

Abstract

This paper presents a model-based approach for stochastic primary balance and public debt simulations to assess fiscal sustainability in selected OECD countries. Fiscal behavior is modeled by means of a fiscal reaction function with time-varying coefficients, which is then, together with a time-varying coefficient vector autoregression, embedded in a stochastic debt sustainability analysis framework. In a pseudo-out-of-sample forecasting exercise using vintage datasets, the model is evaluated against its frequently used fixed coefficient pendant and the European Commission's Economic Forecasts at different horizons. The results indicate that stochastic debt sustainability analyses based on time-varying fiscal reaction functions and vector autoregressions perform competitively in terms of mean squared error and forecast bias at different horizons, especially with respect to public debt as well as short-term primary balance forecasts. Thus, models of this sort should be considered for complementary use at policy institutions, using them together with more "discretionary" approaches to fiscal sustainability analysis.

Keywords: Stochastic debt simulation, fiscal reaction function, time variation, state-space models, MCMC

JEL Codes: E62, H68, C32

1 Introduction

The industrialized world is debt-struck. Both the Global Financial Crisis and the European Sovereign Debt Crisis have put pronounced pressure on many countries' public finances. While an unfavorable demographic transition that will drive many governments' age-related expenditures for decades to come, the recent "Covid-19 crisis" and the corresponding fiscal countermeasures undertaken by governments to stabilize economies around the globe have dimmed the fiscal outlook further. Recently, accelerating inflation dynamics have brought monetary hawks back to the scene, potentially further weighing on debt service costs and thus on the sustainability of public finances.

As a result of these developments, fiscal policy's leeway to achieve policy goals (the *fiscal space*) is severely constrained. Moreover, given a dire public finance outlook, pressure from financial markets might exacerbate the situation, endangering fiscal solvency and further restricting governments' fiscal space, requiring a balancing act between stabilization and sustainability objectives.

Amid those times of elevated fiscal distress, Blanchard et al. (2021) recently argued in favor of rethinking European fiscal rules. Against the frequently proposed reinstallation of those rules, the authors argue that alternative measures of judging fiscal sustainability be superior to the Maastricht criteria, granting more flexibility in uncertain times. At the center of the authors' proposal is the concept of *stochastic debt sustainability analysis* (SDSA), which is used both in academia and at policy institutions, see for example Celasun et al. (2006) or Medeiros (2012).

Employing SDSA to assess the sustainability of public finances has several advantages: As it incorporates fiscal reaction functions (FRFs), SDSA is based upon (past) fiscal behavior, thus providing a less arbitrary way of evaluating fiscal sustainability than more judgement-based approaches like the Maastricht criteria or deterministic DSA. Moreover, by estimating the distributions of macroeconomic shocks of interest and then repeatedly drawing from their joint distribution to ultimately obtain projections of the primary balance and public debt, SDSA neatly incorporates the probabilistic nature of public debt projections (see Everaert and Jansen, 2017 and Medeiros, 2012).

With FRFs being a key ingredient of SDSA, it is crucial they be correctly specified. If, instead, a misspecified FRF is used, the implied debt projections could be (severely) misleading. In a recent paper, Berger et al. (2021) argue in favor of specifying FRFs featuring time-varying coefficients: In particular, they propose modeling the fiscal responsiveness to public debt in a time-varying manner, with the fiscal responsiveness potentially driven by debt thresholds, the macroeconomic environment (encompassing interest rates and growth prospects) or political factors. Equally significant is the specification of vector autoregressions (VARs), which constitute the second major building block of SDSAs of the sort conducted by Medeiros (2012): Consistently with the argument for time-varying coefficient FRFs, the empirical importance of employing time-varying coefficient VARs has been stated by many researchers (see for example Koop and Korobilis, 2013). In this paper, I am building on these findings on the usefulness of time-varying coefficient FRFs and VARs by embedding them in a SDSA and assessing such models' ability in forecasting the short-run development of fiscal variables. To this aim, primary balance and debt forecasts of selected OECD countries are evaluated at various horizons and compared to forecasts of a state-of-the-art fixed coefficient model similar to Medeiros (2012), as used for example by Everaert and Jansen (2017) or Paret (2017). Additionally, the forecast performance is judged by comparison with official forecasts of the European Commission (EC).

The contribution of this paper is twofold: First, a simple public debt projection framework featuring time-varying coefficients is provided, aimed at being used (in this or a more extensive form) at policy institutions, jointly with other approaches already in place. For example, these models could be employed in model-averaging forecast exercises to mitigate the potential performance loss resulting from model uncertainty (see e. g. Moral-Benito, 2015). In this regard, the SDSA framework provided here can be thought of as complementary to existing fiscal forecasting approaches. Second, a vintage data-based forecast assessment framework is provided, allowing for more realistic real-time forecast evaluations than "ex-post" forecasts that use data unknown to the forecaster at the time the forecast is made. Frameworks of this kind can be used in the future to assess the forecasting performance of various (S)DSA models.

My findings suggest that SDSA based on time-varying FRFs in spirit of Berger et al. (2021), combined with a simple public debt projection exercise featuring time-varying coefficient VARs (called the "benchmark model" below), provides competitive primary balance and especially public debt forecasts in terms of mean squared errors (MSEs) for a sample of ten OECD countries. The models employed here outperform a time-invariant ("fixed") coefficient pendant in terms of public debt forecasts at all horizons considered and fare similarly with respect to primary balance projections. Moreover, the benchmark model's forecasts come close to European Commission forecasts for public debt and the primary balance at most horizons. In terms of forecast bias, the EC and the benchmark model perform similarly, but while the EC primary balance nowcasts are biased, the benchmark model's are not. Thus, making use of SDSA with time-varying coefficient FRFs and VARs to nowcast fiscal variables might help overcoming the well-documented bias often found in fiscal projections (see e. g. Frankel, 2011). The above findings are quite robust to changes in the sets of predictors of fixed and time-varying parameters, although excluding the output gap coefficient from the set of time-varying coefficients in the VAR hampers the primary balance forecast performance. Despite the sample of forecast errors being limited, especially at the two-year-horizon, the adequate short-term forecast performance of the benchmark model motivates its use in model-averaging exercises at policy institutions.

The remainder of the paper is structured as follows. Section 2 elaborates on the basics of SDSA and lays out the benchmark SDSA model. In section 3, data, priors and the results are presented. Section 4 concludes.

2 Literature review and model

This section briefly reviews the literature and lays out the state-of-the-art fixed coefficient model as well as the SDSA model featuring time-varying coefficients (the benchmark model).

2.1 SDSA basics

SDSA provides a neat way of assessing the state of governments' public finances and is thus widely used at policy institutions such as the IMF, the European Commission (EC) or the ECB.[1] The groundwork for SDSA has been laid out by Celasun et al. (2006), which more recent studies such as Medeiros (2012), Everaert and Jansen (2017) or Paret (2017) have built upon. The basic idea of these approaches is to forecast public debt by means of a debt accumulation equation:

$$debt_t = \frac{1+i_t}{1+g_t} debt_{t-1} - pb_t, \qquad t = 1, 2, ..., T, \tag{1}$$

where $debt_t$ is the public debt-to-GDP ratio (debt ratio) in period t, i_t is the respective nominal interest rate on the debt outstanding, g_t is the nominal GDP growth rate and pb_t is the primary balance (the government budget balance net of interest payments on the debt outstanding). While the primary balance is typically simulated based on a FRF, the remaining determinants' evolution is captured using forecasts from a VAR containing a set of macroeconomic variables. The joint usage of an FRF and a VAR is motivated by the

[1] A nice overview of a comprehensive DSA framework, as conducted at policy institutions, is provided by Bouabdallah et al. (2017), who elaborate on deterministic DSA and stochastic DSA (as well as on other fiscal sustainability indicators). While the deterministic DSA - as the name suggests - covers a variety of scenarios regarding the future evolution of the determinants of fiscal variables (such as interest rates, inflation and output growth) that are *defined by the researcher/ policy maker*, the stochastic DSA is more agnostic in the sense that it uses a purely data-driven approach to determine the evolution of macroeconomic and fiscal indicators. While discretion and therefore deterministic DSA is certainly helpful for policymakers to gauge a country's fiscal sustainability - especially given the amount of information available at major institutions - this paper intends to make a contribution along the lines of *stochastic* DSA, where discretion plays little to no role. Both approaches, together with other sustainability indicators, can then be combined by the policymaker to make an informed decision about the (future) state of public finances.

low frequency of fiscal decision-making: While (major) budget decisions are often made on a yearly base, it is advisable to employ macroeconomic variables such as real interest rates, GDP or inflation at a higher (quarterly) frequency to “capture the signal” in the variables’ short-run dynamics.

More precisely, SDSA based on Celasun et al. (2006) or Medeiros (2012) is conducted using the following steps:[2]

1. Estimate a FRF and a VAR to obtain estimates of their (reduced-form) coefficients and the distributions of shocks to fiscal and macroeconomic variables.

2. Drawing from the distribution of shocks to macroeconomic variables, feed the VAR forecasts of macroeconomic variables - properly transformed - into the FRF to simulate the primary balance.

3. Use the primary balance forecasts obtained in the previous steps to project the public debt ratio.

4. If applicable: Using this forecast, repeat steps 2 and 3 to obtain primary balance and debt forecasts for horizons $h = 2, 3, ...H$.

5. Repeat these steps R times to obtain distributions for the future paths of the primary balance and debt ratios, where R is a sufficiently high number chosen by the researcher.

2.2 The fiscal reaction function

Clearly, the FRF is a crucial determinant of primary balance and public debt projections of the sort laid out in the previous chapter. If misspecified, inference based on the SDSA framework might be misleading. In spirit of Berger et al. (2021), this paper addresses the specification of FRFs, arguing in favor of a time-varying parameter model. More specifically, consider a standard FRF based on Bohn (1998), such as

[2] Some additional information on the “fixed coefficient approach” can be found in appendix B.4.

$$pb_t = \alpha + debt_{t-1}\beta_1 + X_t\gamma + \epsilon_t \quad (2)$$

where α is a constant, X_t is a set of additional regressors (next to the lagged debt ratio) and ϵ_t is a normally distributed error term. However, as argued in Berger et al. (2021), assuming that the coefficients in α, β_1 and γ are constant over time might be too restrictive: The fiscal reaction to changes in public debt might be altered by various things. Among them, the fiscal responsiveness may depend on the level of debt, as argued in Ghosh et al. (2013). For example, governments might be slow to adjust the primary balance at very low debt ratios, more alert once debt rises and "giving up" on fiscal sustainability at very high debt levels.[3] Additionally, macroeconomic factors such as the growth rate of the economy (among other things by altering the country's "tax generating capacities") or the interest rate on the debt (with higher debt service costs reducing fiscal space) may drive the fiscal reaction to debt.

The fiscal responsiveness to other predictors could be time-varying, too. Assume that the lagged primary balance and a measure of the output gap are contained in X_t above. Then γ - as well as β_1 - may be driven by determinants such as the state of the economy (see for example Égert (2014) on differences in the fiscal responsiveness in up- and downturns), institutional changes (for example the Maastricht criteria or the Fiscal Treaty in the European Union) or changes in the political landscape (for example the political orientation of the government, or so-called electoral business cycles, see e. g. Alesina et al., 1993).

For these reasons, instead of using the specification in (2), I follow Berger et al. (2021) in estimating a FRF of the form

[3] Loosely speaking, a debt-dependent fiscal responsiveness of this sort is what has been called *fiscal fatigue* in the literature (see Ghosh et al., 2013).

$$pb_{it} = \alpha_i + H_{it}\beta_t + X_{it}\gamma + \epsilon_{it}, \tag{3}$$

$$\epsilon_{it} = \mu_t + \rho\epsilon_{i,t-1} + u_{it}, \quad u_{it} \sim N(0, \sigma^2_{u_i}), \quad i = 1, 2, ..., N, \quad t = 1, 2, ..., T, \tag{4}$$

where H_{it} is the predictor matrix corresponding to the time-varying parameters, β_t, and X_{it} corresponds to the fixed parameters, γ. Depending on the specification, H_{it} and X_{it} contain the lagged debt ratio, the lagged primary balance (capturing sluggishness in fiscal policy making) and the output gap. By allowing the parameters corresponding to these three predictors to be time-varying, this FRF constitutes a flexible framework to account for changes in the underlying relationship between the predictors and the primary balance.

Note that, as ultimately any SDSA model should be judged by its forecasting abilities, the final choice of time-varying and fixed parameters will be based on the forecasting performance of the different specifications. Further note that this "specification search" is mostly for illustrative purposes, demonstrating that various models featuring time-varying parameters are capable of producing competitive primary balance and public debt forecasts.

Next to time-varying parameters and estimating a *dynamic* FRF (by adding the lagged primary balance as a predictor), the benchmark specification presented above tackles further specification issues, thus differing from the standard specification presented in (2):

1. Since the inclusion of up to three time-varying coefficients leads to a proliferation of parameters, observations along the cross-sectional dimension are included. That is, by employing a fixed effects panel model and pooling β_t and γ along the cross-sectional dimension, identification of the parameters is facilitated. The coefficients $\alpha_i, i = 1, 2, ..., N$ constitute the country-specific constants and are dealt with using within-group demeaned transformations of the variables.[4]

[4]Employing panel models to estimate FRFs is quite common in the literature (see, among others, Ghosh et al., 2013 or Checherita-Westphal and Žďárek, 2017).

2. Following Ghosh et al. (2013), the model allows for an AR(1) error term, thus accounting for autocorrelation in the residuals not captured by the lagged primary balance term.

3. The model is further enriched by letting the variance of the Gaussian white noise process, u_{it}, be country-specific. This means that the model features another source of cross-country heterogeneity (next to the country-specific effects), accounting for the possibility that the average shock to the primary balance might differ in size between countries.

4. To account for (time-varying) unobserved components, affecting all sample countries, a time-varying component (or time fixed effects) μ_t is included in the error term process.

In what follows, a couple of estimation and specification issues will be elaborated upon.

Endogeneity

Clearly, fiscal policy might have a contemporaneous effect on the business cycle, rendering the output gap potentially endogenous in the FRF, which is why it is commonly instrumented in the literature. I will proceed similarly by running an auxilliary regression of the output gap on the exogenous regressors in (3) and instruments of the output gap (its first two lags, following for example Berger et al., 2021) to obtain a fitted, exogenous pendant of the output gap, which is then used in the estimation algorithm outlined below.[5]

Variable choice

The variable choice employed here is obviously not exhaustive. However, this paper provides a simple framework that can serve as a starting point for future research into SDSA models

[5] More extensive ways to deal with endogenous regressors in a time-varying parameter model are thinkable, see for example Everaert et al., 2017 or Kim and Kim, 2011, where the coefficients of the auxiliary regression are obtained directly from the joint parameter distribution. However, the model presented here serves the main purpose of illustrating that time-varying parameter models in general should be considered in SDSA frameworks. More extensive specifications are left for future work.

based on time-varying FRFs. While one reason for the small set of predictors is parsimony and an attempt to avoid overfitting, the other is data availability: The forecast performance evaluation conducted here is based on an extensive dataset. For example, the inclusion of a ("source-consistent") expenditure gap measure would drastically decrease the sample size, rendering the highly parameterized model (nearly) infeasible. Moreover, candidate predictors would have to be (- again, "source-consistently" -) available for any of the vintages considered, thus further reducing the choice of potential regressors.

Non-centered parameterization

So far, nothing has been said about the exact specification of the time-varying parameters, β_t. A common choice would be to model β_t as a random walk, that is, $\beta_t = \beta_{t-1} + \eta_t$, where η_t is an independent white noise process with variance σ_η^2. However, as σ_η^2 is non-negative, for any prior belief on σ_η^2 unequal to zero, one is enforcing a certain degree of time variation, as for any $\sigma_\eta^2 > 0$, the process β_t would be - governed by a certain degree of time variation. In other words, one would be informative as to whether time variation is present in β_t. Employing a non-centered parameterization provides a neat solution to this problem (see Frühwirth-Schnatter and Wagner, 2010). It is given by:[6]

$$\beta_t = \beta_0 + \sigma_\eta \tilde{\beta}_t, \tag{5}$$

$$\tilde{\beta}_t = \tilde{\beta}_{t-1} + \tilde{\eta}_t, \qquad \tilde{\beta}_0 = 0, \qquad \tilde{\eta}_t \sim N(0,1). \tag{6}$$

By setting a non-informative prior, centered around zero, one is uninformative with respect to the question of whether the respective parameter is governed by time variation or not. Thus, to be as agnostic as possible, the NCP will be used instead of the random walk specification in the estimation algorithm presented in the next subsection. Lastly, note that

[6] Note that the non-centered parameterization (NCP) is simply a reparameterization of the random walk process.

the components of σ_η and $\tilde{\beta}_t$ are only jointly identified. However, as elaborated upon in Frühwirth-Schnatter and Wagner (2010), this can be "solved" by introducing a random sign switch of the components in the estimation routine, which is outlined in the next section and the appendix.

Estimation algorithm for the FRF

In the following, the estimation algorithm for the FRF will be laid out. Note that the system of equations in (3), (4), (5) and (6) can be cast into state-space form. Note that the approach below refers to within-group-demeaned variables to get rid of the country-specific intercepts, $\alpha_i, i = 1, 2, ..., N$. Estimating the model using within-group-demeaned variables has the advantage of reducing the amount of parameters to be estimated, while the coefficients of interest should be equal to the model without demeaning (Frisch-Waugh-Lovell theorem, see e. g. Baltagi, 2013).

The estimation algorithm outlined here draws from Berger et al. (2021) and Blake and Mumtaz (2015).[7] Intuitively, the estimation algorithm approximates intractable joint and marginal parameter distributions by repeatedly drawing the parameters from conditional distributions by means of a Markov-Chain-Monte-Carlo (MCMC) algorithm. For notational convenience, define $\theta \equiv (\beta_0', \sigma_\eta', \gamma')'$, $\tilde{\beta} \equiv (\tilde{\beta}_1, \tilde{\beta}_2, ..., \tilde{\beta}_T)'$, $\sigma_u^2 \equiv (\sigma_{u,1}^2, \sigma_{u,2}^2, ..., \sigma_{u,N}^2)'$, $\mu \equiv (\mu_1, \mu_2, ..., \mu_T)$ and y and χ as the dependent variable and the predictor matrix. The estimation algorithm is conducted using the following steps:

1. Sample the normally distributed coefficients β_0, σ_η and γ conditional on the remaining parameters. That is, draw from $p(\theta|\tilde{\beta}, \rho, \mu, \sigma_u^2, y, \chi)$.

2. Sample the time-varying parameters $\tilde{\beta}$ given the remainder of parameters, that is, draw from $p(\tilde{\beta}|\theta, \rho, \mu, , \sigma_u^2, y, \chi)$. Next, perform a random sign switch for σ_η and $\tilde{\beta}$. That is, randomly multiply both sets of parameters with -1 or 1 with the same probability. Finally, construct β_t from its components.

[7] A detailed version of the algorithm can be found in the appendix.

3. Conditional on the remaining parameters, sample the AR(1) parameter of the error term (ρ) and the unobserved component vector μ, then sample the regression error variances σ_u^2. That is, draw from an independent Normal-inverted Gamma distribution, $p(\rho, \mu, \sigma_u^2 | \theta, \tilde{\beta}, y, \chi)$.

4. Repeat steps 1. to 3. 2*R times and discard the first R draws. If R is a sufficiently high number, the retained R draws provide adequate approximations to the marginal posterior distributions of the parameters.

2.3 BVAR methodology

Following Celasun et al. (2006) and Medeiros (2012), a VAR is used to estimate the correlations between the macroeconomic variables linked to the primary balance and the public debt evolution. Given estimates of these correlations and of the joint distribution of shocks to these variables, one can compute forecasts that can be fed into the primary balance and the debt accumulation equation.

Unlike Celasun et al. (2006) or Medeiros (2012), I employ a VAR that features time-varying slope coefficients, consistent with the time-varying FRF outlined above. Thus, for each country, the VAR model in reduced form can be written as

$$y_t = \phi_{1,t} y_{t-1} + \phi_{2,t} y_{t-2} + ... + \phi_{p,t} y_{t-p} + u_t, \qquad u_t \sim N(0, \Sigma), \tag{7}$$

$$\Phi_t = \Phi_{t-1} + e_t, \qquad e_t \sim N(0, Q), \tag{8}$$

$t = \{1, 2, ..., T_q\}$, where T_q is the number of quarterly observations in the VAR, y_t is a $M \times 1$ vector of demeaned endogenous variables, $\phi_{j,t}$, $j = 1, 2, ..., p$ are $M \times M$ coefficient matrices corresponding to the respective lag matrix y_{t-j} and u_t is a $M \times 1$ vector of reduced-form shocks. The time-varying parameters are collected in $\Phi_t \equiv (vec(\phi_{1,t}), vec(\phi_{2,t}), ..., vec(\phi_{p,t}))'$ and are assumed to follow random walk processes with joint error covariance matrix Q, as outlined in (8).

Notice that, since the VAR is country-specific and the amount of data available for estimating the VAR is restricted, I follow Celasun et al. (2006) and Medeiros (2012) in setting the number of lags in the VAR to two. Parameter proliferation due to time-varying slope coefficients puts further strain on estimation feasibility. To overcome this, a Bayesian VAR (BVAR) is employed: By combining the data with prior information, one can drastically improve upon estimation efficiency. Details on the Bayesian estimation of the VAR are outlined below.[8]

Variable choice

The variable choice for the VAR is broadly in line with Medeiros (2012): Among the variables included are the quarterly growth rate of real GDP, the GDP deflator-based inflation rate and an unweighted average of short-term and long-term real interest rates (see appendix for details). For all countries, the real GDP growth rate and the above-defined average real interest rate for Germany are included (obviously, except for Germany). I deviate from Medeiros (2012) by not including the natural logarithm of the real effective exchange rate, as its inclusion would significantly decrease the sample size.

Estimation algorithm for the BVAR

Analogously to the FRF estimation algorithm outlined above, an MCMC scheme is employed to approximate the posterior distributions of interest. In particular, following Blake and Mumtaz (2015), the algorithm consists of the following steps:[9]

1. Sample the time-varying coefficients Φ_t for $t = 1, 2, ..., T_q$ conditional on the other parameters of the model. That is, draw from p($\Phi_t|\Sigma$,Q,y), using the forward-filtering

[8]While in many applications Σ is allowed to be time-varying (see for example Primiceri, 2005 or Clark and Ravazzolo, 2015), in this model Σ is assumed to be constant over time. This is mainly a practical choice: Adding time-varying volatility to the model drastically increases the number of draws required for adequately approximating the posterior distributions of interest. In fact, it turns out that the number of draws required for convergence is increased so much that running the full SDSA (for *all* vintages) featuring such a VAR model is not feasible given the computing power at my disposal and is thus left for future work.

[9]For more details on the algorithm, see appendix B.2.

backward-sampling algorithm of Carter and Kohn (1994).

2. Sample the state disturbance variance-covariance matrix of the time-varying parameter equation (Q) from its conditional distribution. That is, draw from an inverse Wishart distribution, p($Q|\Phi$,Σ,y), where $\Phi \equiv (\Phi_1', \Phi_2', ..., \Phi_{T_q}')'$.

3. Sample the variance-covariance matrix of the measurement disturbance (Σ) conditional on the other parameters, again from an inverse Wishart distribution. That is, draw from p($\Sigma|\Phi$,Q,y).

4. Repeat steps 1. to 3. 2*R times and discard the first R draws. If R is a sufficiently high number, the retained R draws provide adequate approximations to the marginal posterior distributions of the parameters.

2.4 Simulation of the primary balance and public debt

In this section, the simulation algorithm that repeatedly samples the primary balance and the public debt ratio is laid out. Again, this approach broadly follows Medeiros (2012), but differs at some stages, mainly due to the MCMC algorithms employed for the estimation of the FRF and BVAR coefficients above. The chosen approach will be briefly outlined here. For more details, the reader is referred to the appendix.

Given the parameter estimates of the FRF and the BVAR, the projection algorithm comprises repeatedly drawing future realizations of the macroeconomic variables in the VAR, and then feeding their realized paths into the FRF and the debt accumulation equation. Thus, the algorithm consists of the following steps:

1. Draw shocks to the VAR from their joint distribution, forecast the VAR variables (using equations (7) and (8)) and transform them adequately. That is, convert the forecasts to yearly data and compute yearly GDP growth for the debt accumulation equation and construct an output gap forecast to be fed into the FRF to forecast the

primary balance.[10]

2. Given a sample of T yearly observations, simulate the primary balance for period $T+1$, using equations (3), (4), (5) and (6) and the output gap forecast obtained in the previous step.

3. Feed the $T+1$ forecast for the primary balance, together with the relevant VAR forecasts, into the debt accumulation equation (1) to obtain $debt_{T+1}$.[11]

4. Using the forecast for $debt_{T+1}$, go back to steps 2 and 3. Repeat them for period $T+2$.

5. Save the realizations for the primary balance and the public debt ratio and repeat the above steps R times, where R is the number of retained draws in the MCMC algorithm outlined above. This means that for any retained set of parameter draws in the FRF and the VAR, a path for the primary balance as well as the public debt ratio are obtained. In this way, unlike in the case of Frequentist estimation, the uncertainty surrounding the parameter estimates is directly embedded in the projection exercise.

3 Results

This section covers the data employed for the estimation, the priors as well as the results of the SDSA. Note that, to be as agnostic as possible, in the benchmark model all three explanatory variables of the FRF (that is, the lagged primary balance-to-GDP ratio, the lagged debt ratio and the output gap) are modeled featuring time-varying parameters. Due to the use of the non-centered parameterization, together with the agnostic prior on σ_η (as elaborated upon below), this does not mean that time variation is enforced upon the parameters a priori. Instead, the amount of time variation in the coefficients β_t is governed

[10] The output gap is obtained as the cyclical component of the (one-sided) Hodrick-Prescott filtered output series. As typical for quarterly data, λ is set to 1600.

[11] As outlined in the appendix, I follow Medeiros (2012) in using the implicit interest rate on the debt outstanding as the relevant measure for the nominal interest rate in the debt accumulation equation.

by the data. If the amount of time variation in β_t is limited, its estimated path will simply display little time variation and will not deviate much from its time-invariant component (β_0). Hence, the model with three time-varying parameters will be considered below (before robustness is dealt with).

3.1 Data

In the following, the data used for both FRF and BVAR are outlined. For reasons of consistency, the EC's semi-annual *AMECO Economic Forecast* and the OECD's *Economic Outlook database* vintage datasets are employed from a period spanning from autumn 2014 to spring 2019 (see appendix A for more details on the sample selection). Since the datasets are published twice a year, ten vintages are used in total. For each of the vintages, a sample of ten countries is then used, including Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Japan and the Netherlands. The choice of countries is motivated by the availability of data for the pseudo-real time forecast exercise based on vintage data: All OECD countries for which forecasts of the primary balance and the public debt ratio are available from the primary source used here, that is, from AMECO, were considered candidates for the sample. For all of these OECD countries, where reliable vintage data for all variables of the FRF and the BVAR were available from the below-mentioned sources, are then included in the sample. This leads to a total of ten countries.[12] More details on the data, including the choice of the vintage datasets, are provided in appendix A.

[12]However, note that some data issues remain even for some of the ten sample countries. In particular, there is missing data in two of the OECD vintages: In the "autumn" 2015 vintage, both the nominal GDP series and the GDP deflator series are missing for Belgium, while in the "autumn" 2018 vintage, long-term and short-term interest rates, nominal GDP and the GDP deflator series are missing for Greece. This is dealt with in the following way: Where VAR data are missing, data from the previous vintage are included. This implies that instead of actual observations, for the last two quarterly sample observations in these vintages, forecasts are used instead of observations.

3.2 Priors

In this section, the priors for the Bayesian estimation procedure for the FRF and the BVAR are laid out.

3.2.1 FRF priors

First, the priors employed in the FRF are outlined. Notice that, since some of these priors are derived from sample data, the corresponding prior moments differ (very slightly) between vintages. As such, the priors presented here are exemplary and refer to the final vintage in the sample, that is, the spring 2019 vintage.

Gaussian priors

First, parameters with Gaussian priors are outlined. That is, the respective parameters are – a priori – following a Normal distribution of the sort $N(a_0, A_0)$, where a_0 is the prior mean and A_0 is the prior variance. The normally distributed parameters include the prior on the $m \times 1$ vector β_0 (containing the time-invariant parts of the time-varying parameter processes), the prior on the state error standard deviations σ_η and the prior on the k slope coefficients of the regression (measurement) equation, γ.

Prior statistics for the final vintage are presented in table 1. The table shows the prior means of the respective parameters together with the prior standard deviation and the 5th and 95th percentiles of the implied prior distribution. The prior on the $m \times 1$ vector β_0, which can be interpreted as the coefficient vector of the time-varying parameter processes if no time variation was present in those coefficients, is set with means equal to the (Frequentist) within-group two-stage least squares estimates of the model, where all coefficients are fixed (that is, constant over time).[13] Given the limited sample sizes and thus limited information in the data, each parameter in β_0 is assumed to have a prior variance of 0.01, amounting

[13]Note that these fixed coefficients are simply the estimates of the fixed coefficient model ("fixed model"), presented in appendix B.4.

Table 1: Prior choices for the benchmark specification, final vintage

Gaussian priors

$\sim N(a_0, A_0)$		a_0	$\sqrt{A_0}$	5%	95%
Initial state output gap	$\beta_{0,1}$	0.046	0.1	-0.118	0.211
Initial state lagged primary balance	$\beta_{0,2}$	0.760	0.1	0.596	0.925
Initial state lagged debt	$\beta_{0,3}$	0.010	0.1	-0.155	0.174
Standard deviation state error (output gap)	$\sigma_{\eta,1}$	0	0.32	-0.526	0.526
Standard deviation state error (lagged primary balance)	$\sigma_{\eta,2}$	0	0.32	-0.526	0.526
Standard deviation state error (lagged debt)	$\sigma_{\eta,3}$	0	0.32	-0.526	0.526
Residual autocorrelation parameter	ρ	0	3.2	-5.264	5.264
Time fixed effects	μ	$\underline{0}$	$3.2I$	-5.264	5.264

Inverted Gamma prior

$\sim IG(T\frac{\nu_{0,i}}{2}, T\frac{\nu_{0,i}}{2}\sigma_{0,i}^2)$		$\sigma_{0,i}$	$\nu_{0,i}$	5%	95%
Regression standard deviation, Belgium	$\sigma_{u,1}$	0.251	0.1	0.185	0.405
Regression standard deviation, Germany	$\sigma_{u,2}$	0.304	0.1	0.224	0.491
Regression standard deviation, Ireland	$\sigma_{u,3}$	0.382	0.1	0.281	0.617
Regression standard deviation, Greece	$\sigma_{u,4}$	0.685	0.1	0.504	1.104
Regression standard deviation, France	$\sigma_{u,5}$	0.290	0.1	0.213	0.468
Regression standard deviation, Italy	$\sigma_{u,6}$	0.325	0.1	0.239	0.524
Regression standard deviation, Netherlands	$\sigma_{u,7}$	0.296	0.1	0.218	0.477
Regression standard deviation, Austria	$\sigma_{u,8}$	0.201	0.1	0.148	0.324
Regression standard deviation, Finland	$\sigma_{u,9}$	0.468	0.1	0.344	0.755
Regression standard deviation, Japan	$\sigma_{u,10}$	0.410	0.1	0.302	0.662

Notes: This table summarizes the prior distributions for the final vintage (spring 2019) for the benchmark specification. For the inverted Gamma priors, the prior belief about the standard deviation σ_0 is displayed instead of the corresponding variance parameter as this is easier to interpret. Likewise, for the Gaussian priors, $\sqrt{A_0}$ is reported instead of A_0. For the priors on μ, $\underline{0}$ is a T x 1 vector of zeros, and I is the identity matrix of dimension TxT, with T being the number of time periods in the sample.

to a prior standard deviation of 0.1. Thus, the 90% prior density intervals include a wide range of parameter estimates of the respective parameters found in the literature (see e. g. Checherita-Westphal and Žďárek, 2017 for an extensive overview).

For σ_η, the $m \times 1$ vector of standard deviations of the state disturbances, a prior mean vector with all elements equal to zero is assumed. Thus, time variation is not "forced" upon the parameters a priori. In fact, for a prior mean for σ_η equal to 0, its prior distribution will be unimodal and centered around zero, such that - on average - β_t will remain close to β_0 for all $t = 1, 2, ..., T$ a priori. The prior variances of the vector σ_η are set to 0.1 (implying prior standard deviations of approximately 0.32), which implies quite non-informative prior distributions, where 90% of the innovations to the time-varying components of the time-varying parameters (β_t) lie between -0.526 and 0.526.

In the benchmark model, the parameter vector γ is empty, as the parameters corresponding to the output gap, the lagged primary balance and the lagged debt ratio are assumed to follow time-varying processes (implying that X is empty). Thus, there are no prior moments for γ displayed in table 1. In the robustness section below, where some of the slope parameters are assumed to be time-invariant (and where thus γ is not an empty vector), the prior on γ is the same as the prior on β_0 for the respective component, the reason being that β_0 can be interpreted as the time-invariant component of β_t.

Inverted Gamma priors

The country-specific variances, $\sigma^2_{u,i}$, $i = 1, 2, \ldots, N$, are assumed to follow inverted Gamma distributions. That is, for each i, $\sigma^2_{u,i} \sim IG(c_{0,i}, C_{0,i})$, where the shape parameters are given by $c_{0,i} = \nu_{0,i}/2 * T$ and the scale parameters by $C_{0,i} = c_{0,i} * \sigma^2_{0,i}$, where $\sigma^2_{0,i}$ constitutes the prior belief about the respective regression error variance and $\nu_{0,i}$ the corresponding prior strength. $\sigma^2_{0,i}$ is set to be the regression error variance from country-specific (Frequentist) regressions of the primary balance on its first lag, an (instrumented) output gap, lagged debt and a constant. Table 1 summarizes this information for the sample countries. This implies, for example for Greece, that 90% of the shocks to the primary balance lie between -0.86 and 0.86 percent of GDP.

Random walk components of the time-varying parameter processes

For the random walk components of β_t, that is, $\tilde{\beta}_t$, a forward-filtering backward-sampling algorithm is employed. Thus, its priors are based on the Kalman filter (see appendix for more information on the forward-filtering backward-sampling algorithm).

3.2.2 BVAR priors

This section outlines the priors for the BVAR in equations (7) and (8).[14]

Inverted Wishart priors

Both the variance-covariance matrix of the VAR errors, Σ, and the variance-covariance matrix of the state errors, Q, are assumed to follow inverted Wishart distributions a priori. In particular, as outlined in Blake and Mumtaz (2015), the prior for Σ is given as $p(\Sigma) \sim IW(\Sigma_0, T_{\Sigma_0})$, where Σ_0 is the error variance-covariance matrix of the time-invariant pendant of the VAR in equation (7), estimated with ordinary least squares. The shape parameter T_{Σ_0} is simply the sample size of this VAR. Note that usually the training sample, used to inform the priors, is excluded in the main estimation algorithm. However, given the limited sample size at hand for some vintages and countries, the priors here are informed using the whole sample, without exclusion of some observations in the Gibbs sampling scheme.

The prior for Q is given by $p(Q) \sim IW(Q_0, T_0)$, with the prior scale parameter being defined as $Q_0 = P * T_0 * \tau$. Again, the time-invariant coefficient pendant of the VAR in (7) is used to compute $P = \Sigma_0 \otimes (X'X)^{-1}$, X being the predictor matrix of the VAR. The prior shape parameter T_0 is again the sample size (implying that $T_0 = T_{\Sigma_0} = T_q$). τ is a scaling parameter governing the amount of time variation in the slope coefficients inherent in the prior. Following Blake and Mumtaz (2015), this is set to a very small number of 3.51^{-4}, implying an uninformative prior.

Time-varying slope coefficients

The random walk components collected in Φ are sampled using the Carter and Kohn (1994) forward-filtering backward-sampling algorithm, where the priors of Φ are based on the Kalman filter. For more information on the forward-filtering backward-sampling algorithm,

[14] The prior choices are similar to Blake and Mumtaz (2015), the main exception being that more data is used to inform the priors, as elaborated on below.

the reader is referred to the appendix.

Finally, note that since for each vintage the whole sample is used to inform the prior, the prior moments differ slightly across vintages, just as for the FRF priors outlined above.

3.3 FRF results

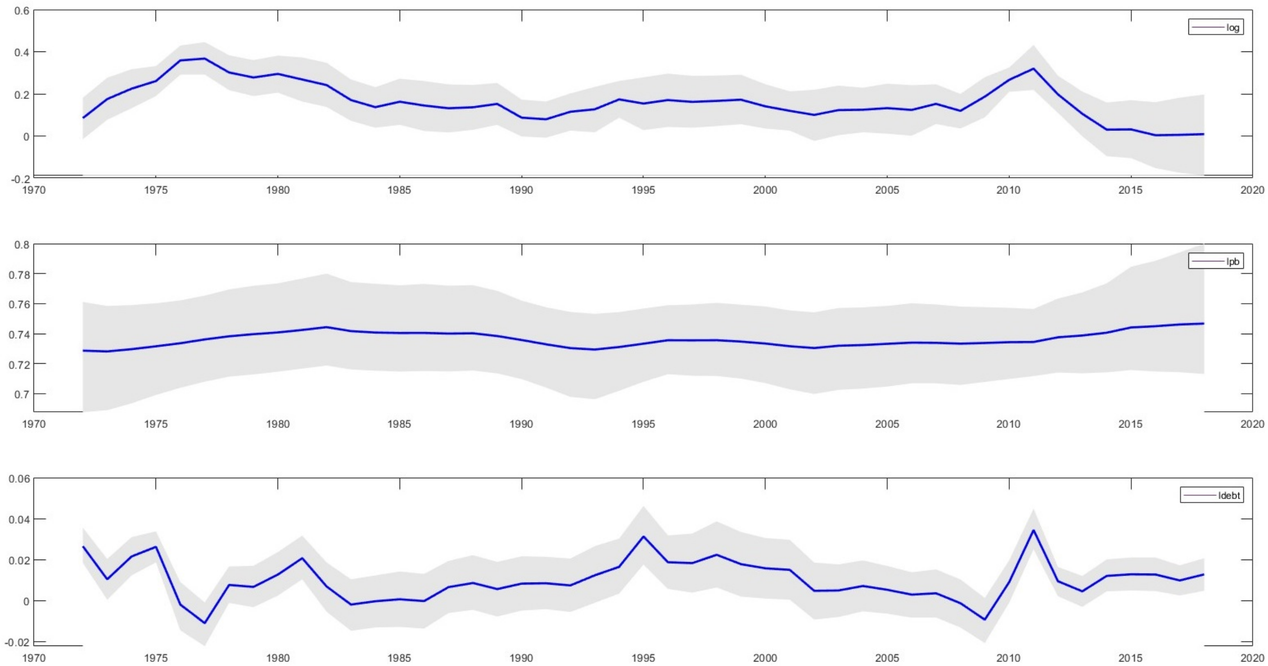
In this section, the FRF results are outlined. Note that for each vintage dataset, the FRF is estimated anew. Given that, due to the similarity of the datasets, the results per vintage are similar and for the sake of clarity, solely the results for the latest vintage (that is, the spring 2019 AMECO vintage) are displayed here, as there the longest available sample is used.

Figure 1 displays the paths of the three time-varying parameters, including their 90% credible sets. The top-most panel displays the evolution of the coefficient on the output gap. First, note that there appears to be a certain degree of time variation present. Such non-linearities in the fiscal response to the business cycle are broadly in line with the literature, in the sense that often an asymmetric response to the cycle in expansions and recessions is modeled (see for example Égert, 2014). Most notably, a pronounced increase in the governments' counter-cyclicality in the aftermath of the Global Financial Crisis is visible: In times of economic distress, stabilizing fiscal measures had been taken to mitigate the effects of the downturn. This effect slowly dissipated over time.

The second panel displays the time-varying parameter linked to the lagged primary balance. This parameter evolves more smoothly (less time-varying) and indicates a high degree of sluggishness in fiscal policy making, again in line with the literature, which argues that it takes time for fiscal policy changes to come about (see Checherita-Westphal and Žďárek, 2017 and Everaert and Jansen, 2018).

The third panel shows the evolution of the time-varying fiscal reaction to public debt. Clearly, the parameter exhibits a substantial degree of time variation, which can be partly explained by the fact that the majority of sample countries are Eurozone members: A decreasing fiscal reaction from the EMU "aspiration period" (i. e. before being granted

Figure 1: Evolution of the time-varying parameters (β_t) in the benchmark specification, final vintage



Notes: The blue lines represent the posterior means of the respective time-varying parameter for the final vintage (spring 2019), with the 90% highest posterior density interval as shaded area. "log" refers to the parameter for the output gap, "lpb" to the lagged primary balance coefficient and "ldebt" to the lagged debt coefficient.

membership to the Eurozone) to the financial crisis is clearly visible. Additionally, the plot shows a pronounced increase in fiscal prudence at the onset of the European Sovereign Debt Crisis.

All said, it is reassuring that the results are broadly in line with the literature, despite the extensive estimation approach and the limited amount of data. However, the overall focus of this paper is to judge the models by the respective forecast performance, which will be done in the next section. Moreover, clearly, changing the set of covariates contained in X and H in (3) will affect in which time-varying parameter paths the variance inherent in the data will show up. Looking at various models in turn, with differing choices of X and H and comparing their forecast performances is advisable (see also the robustness section).

Further results for the benchmark specification are displayed in table 2. Note that in the baseline specification, X and γ in (3) are empty, since the coefficients of the lagged primary

Table 2: Posterior distributions in the benchmark model, final vintage

Sample: 1970-2019, 10 OECD countries

Parameter		**Posterior mean**	**5%**	**95%**
AR(1) parameter of regression error	ρ_1	0.516	0.397	0.632
Measurement error variance, Belgium	σ^2_{u1}	0.068	0.047	0.095
Measurement error variance, Germany	σ^2_{u2}	0.090	0.064	0.123
Measurement error variance, Ireland	σ^2_{u3}	0.233	0.162	0.327
Measurement error variance, Greece	σ^2_{u4}	0.420	0.283	0.603
Measurement error variance, France	σ^2_{u5}	0.057	0.040	0.079
Measurement error variance, Italy	σ^2_{u6}	0.077	0.054	0.107
Measurement error variance, Netherlands	σ^2_{u7}	0.081	0.058	0.111
Measurement error variance, Austria	σ^2_{u8}	0.044	0.030	0.063
Measurement error variance, Finland	σ^2_{u9}	0.230	0.163	0.316
Measurement error variance, Japan	σ^2_{u10}	0.162	0.104	0.245
Implied state error variance, output gap	$\sigma^2_{\eta 1}$	0.0056	0.0021	0.0011
Implied state error variance, primary balance lag	$\sigma^2_{\eta 2}$	0.0001	0.0000	0.0005
Implied state error variance, debt lag	$\sigma^2_{\eta 3}$	0.0002	0.0001	0.0003

Notes: This table summarizes the posterior distributions for the final vintage (spring 2019) for the benchmark specification. The state disturbance variances, σ^2_η, are not estimated directly but implied from the estimate of σ_η in the non-centered parameterization. That is, "Implied state error variance, primary balance lag" is the implied variance of the state disturbance of the time-varying parameter process for the lagged primary balance, and likewise for the fitted output gap and the lagged debt ratio. For reasons of visibility, the nuisance parameters (time fixed effects) are not displayed.

balance, the lagged debt ratio and the (instrumented) output gap are all modeled in a time-varying manner. Thus, table 2 displays only the AR(1) coefficient of the error term, the country-specific variances of the residuals as well as the variances of the state disturbances implied by the estimate for σ_η in the non-centered parameterization.

3.4 SDSA results

In this section, the results of the primary balance and public debt projection exercise are outlined. More precisely, the forecasting performance of the model with respect to forecasting both the primary-balance-to-GDP ratio as well as the public debt ratio are presented. These forecasts are evaluated along two dimensions: The mean squared error (MSE) and the forecast bias from the "true observations" as found in the latest considered vintage of

AMECO data (that is, the spring 2022 vintage). The forecast performance of the benchmark SDSA framework, featuring a time-varying coefficient VAR and a panel time-varying coefficient FRF, is compared to the performance of the fixed coefficient model laid out in section 2.1 and appendix B.4 as well as the European Commission forecast, which has been considered competitive in the past (see Leal et al., 2008).[15] Note that the EC's semi-annual *AMECO Economic Forecast* features forecasts for the current year, one-year-ahead and - for all autumn publications - two-year-ahead point forecasts. Thus, the forecast performance evaluation will be conducted for these horizons. This means that, for any year, two "zero-period-ahead" forecasts (or "nowcasts"), two one-period-ahead forecasts and one two-period-ahead forecasts are made. That is, the forecasts made for 2016 *in 2016*, both in the spring and the autumn vintage, are considered nowcasts below. The forecasts for 2017 made in 2016 are one-period-ahead-forecasts and the 2018 forecast made in autumn 2016 is the two-period-ahead forecast of the year 2016.

In tables 3 - 5, the forecast performance of the model outlined here is compared with those of the EC and the fixed coefficient model. In particular, MSE ratios of the benchmark and the fixed model forecasts against the EC forecasts are displayed, where values smaller than one indicate an advantage of the respective model against the EC. Additionally, the p-values corresponding to the null hypothesis of unbiased forecasts for all three models are presented. Table 3 shows the performances for the zero-period forecast horizon (nowcasts) for both the primary balance and public debt. Clearly, the benchmark model performs better than the fixed model in terms of MSE for both the primary balance and public debt forecasts. When it comes to public debt nowcasts, the benchmark model even outperforms the EC. Additionally, while EC primary balance nowcasts appear to be biased, both the fixed and the benchmark model provide unbiased nowcasts for a significance level of 5%. Regarding

[15]More recent evidence is mixed, with the Commission's performance dependent on the country of interest (see Rybacki et al., 2020). However, their finding that the EC forecasts perform similar to national authorities' forecasts at the horizons considered here still makes the EC projections a valid benchmark for forecast evaluation: If a model's forecast performance comes close to the EC/ national authorities' forecasts, employing it in a sort of model averaging forecast exercise could benefit forecast optimization, as elaborated upon above.

Table 3: Forecast performance evaluation for the primary-balance-to-GDP and the public debt-to-GDP ratios, **nowcasts**

Model	Primary balance		Public debt	
	rMSE	pval (H0: Unbiased)	rMSE	pval (H0: Unbiased)
European Commission	-	0.001	-	0.000
Fixed model	1.823	0.974	0.948	0.000
Benchmark model	1.567	0.086	0.898	0.001

Notes: Presented are Mean Squared Error ratios (rMSEs) of the fixed coefficient model and the benchmark model against the European Commission forecast. Ratios greater than one indicate that the European Commission forecast is superior. Additionally, the table contains p-values for a test of biasedness of forecast errors. That is, the null hypothesis of $\alpha = 0$ in $pb_{it} - pb^F_{itH} = \alpha + u_{itH}$ is tested, where pb_{it} is the actual primary balance in period t for country i and pb^F_{itH} is the corresponding forecast made for period t at period H (similar to An et al., 2018). The results presented here are based on a total of 100 forecast errors.

Table 4: Forecast performance evaluation for the primary-balance-to-GDP and the public debt-to-GDP ratios, **one-period-ahead forecasts**

Model	Primary balance		Public debt	
	rMSE	pval (H0: Unbiased)	rMSE	pval (H0: Unbiased)
European Commission	-	0.265	-	0.201
Fixed model	1.213	0.005	1.229	0.077
Benchmark model	1.279	0.000	1.102	0.380

Notes: Presented are Mean Squared Error ratios (rMSEs) of the fixed coefficient model and the benchmark model against the European Commission forecast. Ratios greater than one indicate that the European Commission forecast is superior. Additionally, the table contains p-values for a test of biasedness of forecast errors. That is, the null hypothesis of $\alpha = 0$ in $pb_{it} - pb^F_{itH} = \alpha + u_{itH}$ is tested, where pb_{it} is the actual primary balance in period t for country i and pb^F_{itH} is the corresponding forecast made for period t at period H (similar to An et al., 2018). The results presented here are based on a total of 100 forecast errors.

the public debt forecasts, all models' nowcasts are biased according to the results based on the sample at hand. However, taken together, these findings clearly motivate the benchmark model's use in model averaging exercises to be conducted at policy institutions, such that pure model-based forecasts like the one presented here can be combined with judgement in order to optimize the fiscal forecast performance.

The above findings are to some extent confirmed for the 1-year-ahead and the 2-year-ahead horizons, as indicated by the results in tables 4 and 5. At both horizons, the benchmark model produces unbiased public debt forecasts, with low MSE ratios against the EC forecast, and clearly outperforming the fixed model. While the benchmark's MSE ratios for the

Table 5: Forecast performance evaluation for the primary-balance-to-GDP and the public debt-to-GDP ratios, **two-period-ahead forecasts**

Model	**Primary balance**		**Public debt**	
	rMSE	**pval (H0: Unbiased)**	**rMSE**	**pval (H0: Unbiased)**
European Commission	-	0.045	-	0.987
Fixed model	1.231	0.006	1.706	0.370
Benchmark model	1.293	0.000	1.361	0.880

Notes: Presented are Mean Squared Error ratios (rMSEs) of the fixed coefficient model and the benchmark model against the European Commission forecast. Ratios greater than one indicate that the European Commission forecast is superior. Additionally, the table contains p-values for a test of biasedness of forecast errors. That is, the null hypothesis of $\alpha = 0$ in $pb_{it} - pb^F_{itH} = \alpha + u_{itH}$ is tested, where pb_{it} is the actual primary balance in period t for country i and pb^F_{itH} is the corresponding forecast made for period t at period H (similar to An et al., 2018). The results presented here are based on a total of 50 forecast errors.

primary balance forecasts are higher (thus worse) than those of the fixed model, the difference is rather small, and with those ratios in the range of 1.2 to 1.3, both models perform quite competitively against the EC. On the downside, the primary balance forecasts of both models are biased, while the EC forecasts are unbiased at least at the one-period-ahead horizon.

Finally note that, given the limited number of vintages, the number of forecast errors to compare is somewhat limited.[16] For each of the nowcasts and one-period-ahead forecasts, 100 forecast errors are given (that is, two forecasts for ten countries per year), while for the two-period-ahead forecasts only 50 forecast errors are available. Thus, especially the two-period-ahead forecasts should be interpreted with caution. Nevertheless, all results point to the forecast performance of the benchmark model being somewhat competitive in relation to both the fixed model and even the EC model.[17]

A formal test that complements the above findings is the Pesaran et al. (2009) panel data version of the Diebold and Mariano (2002) test, which compares the forecasts of two models

[16]The selection of the vintages employed here is based on data consistency reasons, as elaborated upon in the appendix.

[17]Although the models presented here provide biased forecasts for some horizons and variables, this issue might be mitigated in a model averaging exercise. The usage of such a model averaging exercise, featuring FRFs and VARs with time-varying coefficients, is the main proposition of this paper.

of interest.[18] Define the quadratic loss function of a certain variable as

$$z_{it} = \left[e(h)^A_{it}\right]^2 - \left[e(h)^B_{it}\right]^2, \tag{9}$$

$i = 1, 2, ..., N$, $t = 1, 2, ..., T$, where $e(h)^A_{it}$ is the h-period-ahead forecast error for country i in period t for the benchmark model featuring a time-varying coefficient FRF and $e(h)^B_{it}$ is the respective forecast error of the model of comparison, that is, either the fixed coefficient model or the EC forecast. Pesaran et al. (2009) then test the null hypothesis that $\alpha_i = 0$ for all $i = 1, 2, ..., N$ in

$$z_{it} = \alpha_i + \epsilon_{it}, \quad \epsilon_{it} \sim IID(0, \sigma_i^2), \tag{10}$$

the alternative hypothesis being that $\alpha_i < 0$ for some i. The test statistic is computed as

$$\overline{DM} = \frac{\bar{z}}{\sqrt{V(\bar{z})}} \sim N(0, 1), \tag{11}$$

with $\bar{z} \equiv \frac{1}{N}\sum_{i=1}^{N} \bar{z}_i$, $\bar{z}_i \equiv \frac{1}{T}\sum_{t=1}^{T} z_{it}$, $V(\bar{z}) \equiv \frac{1}{NT}\left[\frac{1}{N}\sum_{i=1}^{N} \hat{\sigma}_i^2\right]$, $\hat{\sigma}_i^2 \equiv \frac{\sum_{t=1}^{T}(z_{it}-\bar{z}_i)^2}{T-1}$. For the one-period-ahead and two-period-ahead forecasts ($h = 2$ and $h = 3$), the test statistic is modified to account for autocorrelation in the forecast errors by using a Newey-West type version of $Var(\bar{z}_i)$, see for example Ghysels and Marcellino (2018).

Table 6 displays the results of this test. Values smaller than the 5% critical value of -1.645 indicate a significantly better performance of the benchmark time-varying coefficient model. The strong performance of the benchmark model is confirmed especially by the results against the fixed model, where the DM statistic provides formal evidence for the superiority of the benchmark model in terms of MSE for the nowcasts as well as at the two-period-ahead horizon. At the same time, the benchmark's debt forecast is not outperformed by the EC at any forecast horizon.

[18] The following remarks closely follow Pesaran et al. (2009). For simplicity, whenever it does not contradict the notation used so far, their notation is used.

Table 6: Diebold-Mariano panel test results

Model	**Primary balance**			**Public debt**		
	0p	**1p**	**2p**	**0p**	**1p**	**2p**
European Commission	1.430	2.148	1.205	-0.414	0.986	1.520
Fixed model	-1.729	1.491	1.516	-1.716	-1.093	-1.656

Notes: This table presents the results of the Pesaran et al. (2009) panel data version of the Diebold-Mariano test, where the benchmark model featuring time-varying coefficients is tested against the European Commission forecast and the forecast of the fixed coefficient model. “0p” is the nowcast, “1p” the one-year-ahead and “2p” the two-year-ahead forecasts, respectively. The test is a one-sided test with the null hypothesis that the forecasts from the two models are not significantly different, the alternative hypothesis being that the benchmark model’s forecasts are significantly better. The 5% critical value is -1.645. Thus, values smaller than -1.645 indicate superiority of the benchmark model’s forecasts at the respective horizon.

Additionally, the table shows that the benchmark’s primary balance nowcasts are significantly better than those of the fixed model, while once again the EC forecasts do not have a clear edge over the benchmark model. However, the primary balance forecast performance at the one- and two-period horizon is worse, with the EC forecast’s superiority over the benchmark model even being statistically significant for the one-period horizon. Nevertheless, the DM test results clearly show that a model averaging forecast approach that encompasses an SDSA model that features a time-varying coefficient FRF and VAR might be a helpful contributor to overall fiscal forecasting performance.

3.5 Robustness

In this section, summarized results for two alternative specifications are presented. The second specification is motivated by the fact that the time-varying parameter of the lagged primary balance displays little time variation (see figure 1). In this specification, the coefficient for the lagged primary balance is included as a time-invariant parameter. That is, the lagged primary balance is included as a regressor in X with the corresponding coefficient being included in γ (see equation (3)). The third specification follows the baseline specification in Berger et al. (2021), who find formal evidence for time variation in the lagged debt parameter. Thus, in this specification, only the lagged debt ratio is contained in the matrix

H, while the output gap as well as the lagged primary balance are included in X. Thus, their parameters are treated as fixed (in the sense of not time-varying) in this specification.

Table 7 illustrates the forecast performance of all three specifications for all three forecast horizons. The table also repeats the results of the fixed model for reasons of comparability. Clearly, the differences in forecast performance between the benchmark specification and specification 2 are negligible. While the competitive public debt forecast performance is also visible for specification 3, its primary balance forecasts are somewhat worse, especially with respect to the nowcasts. Given the amount of time variation of the output gap coefficient in the benchmark specification, displayed in figure 1, the poor forecast results might be seen as a preliminary indication of model misspecification stemming from forcing the output gap coefficient to be time-invariant in specification 3. However, as indicated by the Pesaran et al. (2009) test results in table 8, the fixed model still does not (significantly) outperform the benchmark model in terms of primary balance forecast at any horizon. Thus, even specification 3 might be worth considering in a model averaging forecast exercise, especially due to its strong public debt forecast performance.

Taken together, these findings provide some evidence that simple SDSA models featuring time-varying coefficient FRFs and VARs deserve some praise when it comes to fiscal forecasting. This finding is robust to changes in the specification, especially when it comes to the public debt forecast performance.

Table 7: Forecast performance evaluation for the primary-balance-to-GDP and the public debt-to-GDP ratios, **all horizons, alternative specifications**

- Nowcasts -

Model	Primary balance		Public debt	
	rMSE	pval (H0: Unbiased)	rMSE	pval (H0: Unbiased)
Fixed model	1.823	0.974	0.948	0.000
Benchmark model	1.567	0.086	0.898	0.001
Specification 2	1.558	0.062	0.894	0.001
Specification 3	2.314	0.000	0.833	0.009

- One-period-ahead forecasts -

Model	Primary balance		Public debt	
	rMSE	pval (H0: Unbiased)	rMSE	pval (H0: Unbiased)
Fixed model	1.213	0.005	1.229	0.077
Benchmark model	1.279	0.000	1.102	0.380
Specification 2	1.286	0.000	1.100	0.404
Specification 3	1.358	0.000	0.988	0.820

- Two-period-ahead forecasts -

Model	Primary balance		Public debt	
	rMSE	pval (H0: Unbiased)	rMSE	pval (H0: Unbiased)
Fixed model	1.231	0.006	1.706	0.370
Benchmark model	1.293	0.000	1.361	0.880
Specification 2	1.289	0.000	1.356	0.902
Specification 3	1.589	0.000	1.275	0.391

Notes: Presented are Mean Squared Error ratios (rMSEs) of the fixed coefficient model, the benchmark model as well as two further specifications of the benchmark model (specifications 2 and 3) against the European Commission forecast. Ratios greater than one indicate that the European Commission forecast is superior. Additionally, the table contains p-values for a test of biasedness of forecast errors. That is, the null hypothesis of $\alpha = 0$ in $pb_{it} - pb_{it}^F = \alpha + u_{itH}$ is tested, where pb_{itH} is the actual primary balance in period t for country i and pb_{itH}^F is the corresponding forecast made for period t at period H (similar to An et al., 2018). The results are based on 100 forecast errors for the nowcast and one-period-ahead horizon and 50 forecast errors for the two-period-ahead horizon.

Table 8: Diebold-Mariano panel test results

Model	**Primary balance**			**Public debt**		
	0p	**1p**	**2p**	**0p**	**1p**	**2p**
Benchmark vs. EC	1.430	2.148	1.205	-0.414	0.986	1.520
Benchmark vs. fixed model	-1.729	1.491	1.516	-1.716	-1.093	-1.656
Specification 2 vs. EC	1.365	2.168	1.319	-0.452	0.988	1.524
Specification 2 vs. fixed model	-1.717	1.520	1.447	-1.790	-1.112	-1.652
Specification 3 vs. EC	1.738	1.502	2.278	-1.505	0.672	1.662
Specification 3 vs. fixed model	-0.264	0.315	1.527	-1.907	-1.251	-1.567

Notes: This table presents the results of the Pesaran et al. (2009) panel data version of the Diebold-Mariano test, where specifications 2 and 3 are tested against the European Commission forecast and the forecast of the fixed coefficient model. "0p" is the nowcast, "1p" the one-year-ahead and "2p" the two-year-ahead forecasts, respectively. The test is a one-sided test with the null hypothesis that the forecasts from the two models are not significantly different, the alternative hypothesis being that the benchmark model's forecasts are significantly better. The 5% critical value is -1.645. Thus, values smaller than -1.645 indicate superiority of the benchmark model's forecasts at the respective horizon.

4 Conclusion

In times of Covid-19 and the corresponding countermeasures, taken by governments around the globe to stabilize struggling economies, questions of public debt sustainability are as relevant as ever. This article looks at fiscal sustainability in spirit of Blanchard et al. (2021), who argue in favor of a rethinking of European fiscal rules, with stochastic debt sustainability analysis (SDSA) playing a key role in their proposal.

The above findings suggest that SDSAs based on time-varying fiscal reaction functions in spirit of Berger et al. (2021) and time-varying coefficient vector autoregressions, combined with a simple public debt projection exercise as in Medeiros (2012), provide competitive primary balance and especially public debt forecasts in terms of mean squared errors (MSEs). The benchmark model outperforms a time-invariant coefficient pendant with respect to public debt forecasts and additionally fares better in terms of primary balance nowcasts. Moreover, the mostly low MSE ratios in comparison to the European Commission forecast indicate a considerable forecast precision of the benchmark model. In terms of forecast bias, the models often perform similarly, but EC primary balance nowcasts are biased, while the benchmark model delivers unbiased forecasts. Thus, when used complementary, the benchmark model might be a contributor to tackling the well-documented fiscal forecast bias (see e. g. Frankel, 2011).

Given the adequate forecast performance of the SDSA frameworks featuring time-varying FRFs and VARs presented here, I argue that such models should be considered for usage in broader DSA frameworks as conducted at policy institutions, for example by means of model averaging exercises to mitigate the potential performance loss resulting from model uncertainty (see e. g. Moral-Benito, 2015).

Time-varying FRFs can and should be used in model-averaging forecast exercises at policy institutions. However, extensions to the simple illustrative model presented here could be considered. First, the set of covariates employed in the FRF (as well as the VAR) is not extensive, which is partly owed to data limitations that occurred due to the usage of vintage datasets for the pseudo-out-of-sample forecast evaluation. If data issues might (at least for some countries) be resolved, one might consider using more predictors (for example an expenditure gap as in Bohn, 1998). Moreover, looking into alternative forms of non-linearities in the FRF, such as regime-switching rules (see for example Legrenzi and Milas, 2013) might be fruitful. Further aspects to be investigated are the handling of endogenous regressors (see e. g. Kim and Kim, 2011) or accounting for the feedback link of fiscal policy on the macroeconomy (see e. g. Everaert and Jansen, 2017). Regarding the VAR specification, one might consider incorporating time-varying volatility parameters as in Primiceri (2005) or Clark and Ravazzolo (2015), especially if equipped with the necessary computing power.

Similar to what has been done here, one might evaluate the quality of alternative models through the lens of their forecasting performances, using fiscal vintage data. Given the results at hand, this might certainly be worth considering. In spirit of Blanchard et al. (2021), one might come up with a full-fledged framework to complement fiscal sustainability measures currently in place.

References

Alesina, A., Cohen, G. D., and Roubini, N. (1993). Electoral business cycle in industrial democracies. *European Journal of Political Economy*, 9(1):1–23.

An, Z., Jalles, J. T., Loungani, P., and Sousa, R. M. (2018). Do imf fiscal forecasts add value? *Journal of Forecasting*, 37(6):650–665.

Baltagi, B. (2013). *Econometric analysis of panel data*. John Wiley & Sons.

Berger, T., Dubbert, T., and Schoonackers, R. (2021). Fiscal prudence: It's all in the timing–estimating time-varying fiscal policy reaction functions for core eu countries. *Available at SSRN 3810841*.

Blake, A. and Mumtaz, H. (2015). *Applied Bayesian Econometrics for central bankers*. Number 36 in Handbooks. Centre for Central Banking Studies, Bank of England.

Blanchard, O., Leandro, A., and Zettelmeyer, J. (2021). Redesigning EU fiscal rules: from rules to standards. *Economic Policy*, 36(106):195–236.

Bohn, H. (1998). The Behavior of U. S. Public Debt and Deficits*. *The Quarterly Journal of Economics*, 113(3):949–963.

Bouabdallah, O., Checherita-Westphal, C. D., Warmedinger, T., De Stefani, R., Drudi, F., Setzer, R., and Westphal, A. (2017). Debt sustainability analysis for euro area sovereigns: a methodological framework. *ECB Occasional Paper*, (185).

Carter, C. K. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553.

Celasun, O., Ostry, J. D., and Debrun, X. (2006). Primary surplus behavior and risks to fiscal sustainability in emerging market countries: A "fan-chart" approach. *IMF Staff Papers*, 53(3):401–425.

Checherita-Westphal, C. and Žďárek, V. (2017). Fiscal reaction function and fiscal fatigue: Evidence for the euro area. ECB Working Paper 2036, Frankfurt a. M.

Clark, T. E. and Ravazzolo, F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, 30(4):551–575.

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 20(1):134–144.

Durbin, J. and Koopman, S. J. (2012). *Time series analysis by state space methods*. Oxford university press.

Égert, B. (2014). Fiscal policy reaction to the cycle in the oecd: pro-or counter-cyclical? *Mondes en developpement*, (3):35–52.

Everaert, G. and Jansen, S. (2017). On the estimation of panel fiscal reaction functions: Heterogeneity or fiscal fatigue? national bank of belgium working paper no. 320.

Everaert, G. and Jansen, S. (2018). On the estimation of panel fiscal reaction functions: Heterogeneity or fiscal fatigue? *Economic Modelling*, 70:87 – 96.

Everaert, G., Pozzi, L., and Schoonackers, R. (2017). On the stability of the excess sensitivity of aggregate consumption growth in the usa. *Journal of Applied Econometrics*, 32(4):819–840.

Frankel, J. (2011). Over-optimism in forecasts by official budget agencies and its implications. *Oxford Review of Economic Policy*, 27(4):536–562.

Frühwirth-Schnatter, S. and Wagner, H. (2010). Stochastic model specification search for gaussian and partial non-gaussian state space models. *Journal of Econometrics*, 154(1):85–100.

Ghosh, A. R., Kim, J. I., Mendoza, E. G., Ostry, J. D., and Qureshi, M. S. (2013). Fiscal Fatigue, Fiscal Space and Debt Sustainability in Advanced Economies. *The Economic Journal*, 123(566):F4–F30.

Ghysels, E. and Marcellino, M. (2018). *Applied economic forecasting using time series methods*. Oxford University Press.

Kim, Y. and Kim, C. (2011). Dealing with endogeneity in a time-varying parameter model: joint estimation and two-step estimation procedures. *The Econometrics Journal*, 14(3):487–497.

Koop, G. and Korobilis, D. (2013). Large time-varying parameter vars. *Journal of Econometrics*, 177(2):185–198. Dynamic Econometric Modeling and Forecasting.

Leal, T., Pérez, J. J., Tujula, M., and Vidal, J.-P. (2008). Fiscal forecasting: Lessons from the literature and challenges*. *Fiscal Studies*, 29(3):347–386.

Legrenzi, G. and Milas, C. (2013). Modelling the fiscal reaction functions of the gips based on state-varying thresholds. *Economics Letters*, 121(3):384–389.

Mauro, P., Romeu, R., Binder, A., and Zaman, A. (2015). A modern history of fiscal prudence and profligacy. *Journal of Monetary Economics*, 76:55–70.

Medeiros, J. (2012). Stochastic debt simulation using VAR models and a panel fiscal reaction function – results for a selected number of countries. European Economy - Economic Papers 2008 - 2015 459, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Moral-Benito, E. (2015). Model averaging in economics: An overview. *Journal of Economic Surveys*, 29(1):46–75.

Paret, A.-C. (2017). Debt sustainability in emerging market countries: Some policy guidelines from a fan-chart approach. *Economic Modelling*, 63:26–45.

Pesaran, M. H., Schuermann, T., and Smith, L. V. (2009). Forecasting economic and financial variables with global vars. *International Journal of Forecasting*, 25(4):642–675. Special section: Decision making and planning under low levels of predictability.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy. *The Review of Economic Studies*, 72(3):821–852.

Rybacki, J. et al. (2020). Are the european commission's forecasts of public finances better than those of national governments? *Central European Economic Journal*, 7(54):101–109.

A Data

This section provides details about the data used in this paper. Tables A.9 summarizes the information for the data used for the FRF. Note that for the fiscal reaction function, for the primary balance-to-GDP ratio and the debt-to-GDP ratio, AMECO data are used as long as available. For those countries where fiscal AMECO data are not dating back all the way to the beginning of the sample - that is, to 1970 - the respective series is complemented using data from Mauro et al. (2015), using retropolation as in Berger et al. (2021). Thereby, a higher number of observations along the time dimension is obtained, with the panel dataset being balanced, ensuring that at any point in time, the degree of time variation in the time-varying parameters is driven by all countries jointly and not only by a subgroup of them. Implicit interest rates and stock-flow adjustments, obtained from AMECO as well, are used in the debt accumulation equation, as elaborated upon below.

For the VAR part, quarterly data are employed to capture correlations between the variables of interest that are more frequent than the yearly frequency for AMECO data. The variables used in the VAR and its sources are summarized in table A.10. Further note that the data limitations faced in the VAR part differ between countries. For each VAR, the longest sample available is used. Country-specific data availabilities are summarized in table A.11.

Handling of the vintages and data issues

To assess the pseudo-real time forecasting performance of primary balance and public debt projections, ten vintages with yearly data are used. The choice of vintages is motivated by reasons of consistency: The first vintage used is the AMECO dataset from autumn 2014, being the first dataset based on the European system of accounts (ESA) 2010. Using vintages before that would be problematic especially with respect to the output gap variable, as the change in accounting standards implied major revisions in the series. Thus, all vintages based on ESA 2010 standards are used for the forecasting performance evaluation to ensure a high degree of consistency between datasets. As "true values", primary balance and public debt ratios using the latest available vintage, i. e. the spring 2022 vintage, are used. This implies that the "true values" for the periods for which the latest forecasts are made (that is, 2019-2020 in the spring 2019 vintage) have all been subject to at least two revisions.

In order to realistically assess the forecasting performance of the SDSA - to avoid hindsight bias - at the moment of forecasting, only the data already available to the forecaster can be used. This implies two things:

Table A.9: Data description for the fiscal reaction function and the debt accumulation equation

Series name	Sources	Transformation
Primary balance	Mauro et al. (2015), AMECO's "Net lending (+) or net borrowing (-) excluding interest: general government :- Excessive deficit procedure"	Percentages of GDP
Public debt	Mauro et al. (2015), AMECO's "General government consolidated gross debt :- Excessive deficit procedure (based on ESA 2010) and former definitions (linked series)"	Percentages of GDP
Output gap	AMECO's "Gap between actual and potential gross domestic product at 2010 reference levels"	Percentages of potential GDP
Implicit interest rate	AMECO's "Implicit interest rate: general government :- Interest as percent of gross public debt of preceding year Excessive deficit procedure (based on ESA 2010)"	Percentage of gross public debt
Stock-flow adjustments	AMECO's "Stock-flow adjustment on general government consolidated gross debt :- Excessive deficit procedure (based on ESA 2010) "	Percentages of GDP

Table A.10: Data description for the vector autoregression

Series name	Sources	Transformation
Real Gross Domestic Product (GDP)	OECD Economic Outlook database series "Gross domestic product, nominal value, market prices", deflated by "Gross domestic product, market prices, deflator"	Δln
Real interest rate	Unweighted average of the OECD Economic Outlook database series "Long-term interest rate on government bonds" and "short-term interest rate", adjusted for year-on-year inflation using the GDP deflator	-
Inflation	OECD Economic Outlook database series "Gross domestic product, market prices, deflator"	-

Table A.11: Data availability in the VAR

Country	Earliest data availability
Austria	1970 Q1
Belgium	1960 Q1
Finland	1970 Q1
France	1970 Q1
Germany	1991 Q1
Greece	1995 Q1
Ireland	1990 Q1
Italy	1971 Q1
Japan	1969 Q1
Netherlands	1960 Q1

1. For the data taken from AMECO, at each point in time, the respective vintage publishing the EC forecasts is used.

2. For the OECD data (used for the VAR), the latest vintage available at the moment the EC vintage is published, is used.

There is one restriction to the second rule: While the AMECO vintages are always published in May (spring release) and November (autumn release) of the respective year, the OECD vintage publication date varies slightly from year to year for the respective releases. For example, most of the time, when the AMECO spring vintage is released, the OECD vintage containing information up to the *first quarter* of the respective year is available. However, in some cases, the OECD release occurs after the AMECO release date. If that is the case, technically, information (for one or two quarterly observations) is used that would not be available to the forecaster the moment the forecast is made, implying a slight information advantage for the forecasts made here. However, this advantage is small and is still a major improvement over systematically using ex-post data such as the latest data available. Given that very few observations in the sample are concerned, this circumstance is ignored for simplicity.

There is another data-related issue concerning the OECD vintages: In the "autumn" 2015 OECD vintage, both the nominal GDP series and the GDP deflator series are missing for Belgium, while in the "autumn" 2018 vintage, long-term and short-term interest rates, nominal GDP and the GDP deflator series are missing for Greece. This is dealt with in the following way: Where VAR data are missing, data from the previous vintage are included. This implies that instead of actual observations, for the last two quarterly sample observations (only), forecasts are used instead. Again, only few observations are affected.

B Stochastic debt sustainability analysis algorithm

This section lays out the complete stochastic debt simulation analysis employed here. The procedure will be outlined in three subsections, dealing with the empirical FRF, the BVAR and the fiscal projection algorithm in turn.

B.1 Fiscal reaction function

In this section, the Gibbs sampling algorithm, used to estimate the coefficients of the time-varying panel FRF, is laid out. The full model consists of the equations (3), (4), (5) and (6), restated here for convenience (with slight notational differences, as elaborated upon below):

$$pb_{it} = H_{it}\beta_t + X_{it}\gamma + \epsilon_{it}, \tag{B.1}$$

$$\epsilon_{it} = \mu_t + \rho\epsilon_{i,t-1} + u_{it} \qquad u_{it} \sim N(0, \sigma^2_{u_i}), \tag{B.2}$$

$$\beta_t = \beta_0 + \sigma_\eta \tilde{\beta}_t, \tag{B.3}$$

$$\tilde{\beta}_t = \tilde{\beta}_{t-1} + \tilde{\eta}_t, \qquad \tilde{\beta}_0 = 0, \qquad \tilde{\eta}_t \sim N(0, 1), \tag{B.4}$$

where $i = 1, 2, ..., N$, $t = 1, 2, ..., T$, pb_{it} is the primary balance, H_{it} is the matrix of predictors corresponding to the $m \times 1$ vector of time-varying parameters, β_t, X_{it} is the predictor matrix corresponding to the coefficients that are assumed to be fixed (γ). Note that *all variables are within-group demeaned*. For simplicity and for reasons of parsimony, the demeaning as well as the auxilliary regression to account for the endogeneity of the output gap, elaborated upon above, are conducted prior to the Markov-Chain-Monte-Carlo algorithm presented here. Following, among others, (Ghosh et al., 2013), some persistence (autocorrelation of order 1) is accounted for in the regression (measurement) error (ϵ_{it}). Additionally, time-varying unobserved components (time fixed effects) are accounted for by including μ_t.

(B.3) and (B.4) constitute a non-centered parameterization (NCP) of the time-varying parameters (see e. g. Frühwirth-Schnatter and Wagner, 2010). While a simple random walk parameterization of the time-varying parameters would "force" the parameters into a time-varying direction for any state error disturbance with variance greater zero (see e. g. Berger et al., 2021), this parameterization has the advantage that it is quite agnostic as to whether time variation is present in the data. This is the case since σ_η is assumed to be normally distributed in the NCP, with an assumed prior mean equal to zero. Thus, if the data informs β_t to be constant for $t = 1, 2, ...T$, the β_t based on the NCP will not wander off significantly from β_0.

In what follows, details on the MCMC algorithm to jointly sample the time-varying parameter vectors in β, the hyperparameters β_0, σ_η, γ, μ, ρ and σ_u^2 are provided. This section draws from Berger et al. (2021).

B.1.1 Sampling the parameters β_0, σ_η and γ

In this block, the regression parameters β_0, σ_η and γ are sampled conditionally on the time-varying parameters (β_t), the AR(1) coefficient of the autocorrelated error terms (ρ), the time fixed effects (μ_t) and the country-specific regression error variances, collected in σ_u^2. For notational convenience, define a general regression model

$$y = \chi\theta + e, \quad e \sim N(0, \Sigma), \tag{B.5}$$

where y is the dependent variable vector and χ is a predictor matrix corresponding to the parameter vector $\theta \equiv (\beta_0', \sigma_\eta', \gamma')'$. For both y and χ, observations are stacked over cross-sectional and time units, that is, over $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$, with i being the slower index. The covariance matrix of the error term e is a diagonal matrix given by $\Sigma = diag\,(\sigma_u^2 \otimes \iota_T)$, where σ_u^2 is the $N \times 1$ vector of country-specific variances ($\sigma_u^2 \equiv (\sigma_{u,1}^2, \sigma_{u,2}^2, ...\sigma_{u,N}^2)'$) and ι_T is a $T \times 1$ vector of ones.

A Normal prior with $\theta \sim N(a_0, A_0)$ is assumed, where a_0 is the vector of prior means of the respective parameters and A_0 is the prior variance-covariance matrix. As this prior is conjugate, it implies a normally distributed posterior, that is, $p(\theta|\tilde{\beta}, \mu, \rho, \sigma_u^2, y, \chi) \sim N(a_T, A_T)$, where

$$a_T = A_T \left(\chi'\Sigma^{-1}y + A_0^{-1}a_0\right), \tag{B.6}$$

$$A_T = \left(\chi'\Sigma^{-1}\chi + A_0^{-1}\right)^{-1}. \tag{B.7}$$

The above can then be applied to the state-space model in equations (B.1)-(B.4): First, transform the measurement equation such that its error terms are white noise. That is, insert (B.2) into (B.1) and rewrite to obtain:

$$pb_{it}^* = H_{it}^*\beta_t + X_{it}^*\gamma + u_{it}, \tag{B.8}$$

where $pb_{it}^* = pb_{it} - \mu_t - \rho pb_{i,t-1}$ and analogously for H_{it}^* and X_{it}^*. Note that the errors in the transformed model, $u_{it} = \epsilon_{it} - \mu_t - \rho\epsilon_{i,t-1}$ are normally distributed. Next, inserting (B.3) into (B.8) yields

$$pb_{it}^{*} = H_{it}^{*}\beta_0 + H_{it}^{*}\sigma_\eta\tilde{\beta}_t + X_{it}^{*}\gamma + u_{it}, \quad u_{it} \sim N(0, \sigma_{u_i}^2), \tag{B.9}$$

which can be written as

$$\underbrace{pb_{it}^{*}}_{y_{it}} = \underbrace{\begin{bmatrix} H_{i,t}^{*} & H_{i,t}^{*}\tilde{\beta}_t & X_{it}^{*} \end{bmatrix}}_{\chi_{it}} \underbrace{\begin{bmatrix} \beta_0 \\ \sigma_\eta \\ \gamma \end{bmatrix}}_{\theta} + u_{it}. \tag{B.10}$$

θ can then be sampled from $p(\theta|\tilde{\beta}, \mu, \rho, \sigma_u^2, y, \chi) \sim N(a_T, A_T)$, where the posterior moments are given by (B.6) and (B.7).

B.1.2 Sampling the time-varying parameters

In this block, the forward-filtering backward-sampling procedure of Carter and Kohn (1994) is employed to sample the time-varying component $\tilde{\beta}$ given θ, μ, ρ and σ_u^2. The conditional linear Gaussian state-space model is given by

$$y_t = H_t s_t + e_t, \qquad e_t \sim MN(0_N, R), \tag{B.11}$$

$$s_t = F s_{t-1} + K_t v_t, \qquad s_0 \sim N(b_0, V_0), \qquad v_t \sim N(0, Q), \tag{B.12}$$

where y_t is an N x 1 vector of observations and H_t is the predictor matrix, with s_t being the corresponding time-varying parameter vector. The matrices χ, F, K, R, Q as well as the expected value and variance of the initial state s_0, that is, b_0 and P_0, are assumed to be known (conditioned upon). The disturbances e_t and v_t are assumed to be serially uncorrelated and independent of each other for $t = 1, 2, ..., T$. For details on the linear Gaussian state-space model, see Durbin and Koopman (2012).

The Kalman filter can then be employed on this linear Gaussian state-space model to filter the unknown state s_t (forward-filtering). s_t can then be sampled from its conditional distribution (backward-sampling), as described in Carter and Kohn (1994).

Rearrange terms in equation (B.9) to obtain, together with the state equation (B.4), the conditional state-space model for $\tilde{\beta}_t$:

$$\overbrace{pb^*_{it} - H^*_{it}\beta_0 - X^*_{it}\gamma}^{y_{it}} = \overbrace{H^*_{it}\sigma_\eta}^{H_t} \overbrace{\tilde{\beta}_t}^{s_t} + \overbrace{u_{it}}^{e_{it}}, \qquad u_{it} \sim N(0, \overbrace{diag(\sigma^2_u)}^{R}), \tag{B.13}$$

$$\underbrace{\tilde{\beta}_t}_{s_t} = \underbrace{I_m}_{F} \underbrace{\tilde{\beta}_{t-1}}_{s_{t-1}} + \underbrace{I_m}_{K_t} \underbrace{\tilde{\eta}_t}_{v_t}, \qquad \tilde{\eta}_t \sim N(0, \underbrace{I_m}_{Q}), \tag{B.14}$$

where I_m is the identity matrix of dimension m, m being the number of time-varying parameters in the model. Note that the $1 \times m$ vector of states, s_t, is assumed to be homogeneous across countries for each $j = 1, ..., m$. Stacking observations over $i = 1, 2, ..., N$, this can be written as

$$\overbrace{\begin{bmatrix} pb^*_{1t} - H^*_{1t}\beta_0 - X^*_{1t}\gamma \\ \vdots \\ pb^*_{Nt} - H^*_{Nt}\beta_0 - X^*_{Nt}\gamma \end{bmatrix}}^{y_t} = \overbrace{\begin{bmatrix} H^*_{1t}\sigma_\eta \\ \vdots \\ H^*_{Nt}\sigma_\eta \end{bmatrix}}^{H_t} \overbrace{\begin{bmatrix} \tilde{\beta}^1_t \\ \vdots \\ \tilde{\beta}^m_t \end{bmatrix}}^{s_t} + \overbrace{\begin{bmatrix} u_{1t} \\ \vdots \\ u_{Nt} \end{bmatrix}}^{e_t}, \tag{B.15}$$

$$\overbrace{\begin{bmatrix} u_{1t} \\ \vdots \\ u_{Nt} \end{bmatrix}}^{e_t} \sim \left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \overbrace{\begin{bmatrix} \sigma^2_{u_1} & & \\ & \ddots & \\ & & \sigma^2_{u_N} \end{bmatrix}}^{R} \right) \tag{B.16}$$

$$\underbrace{\tilde{\beta}_t}_{s_t} = \underbrace{I_m}_{F} \underbrace{\tilde{\beta}_{t-1}}_{s_{t-1}} + \underbrace{I_m}_{K_t} \underbrace{\tilde{\eta}_t}_{v_t}, \tag{B.17}$$

$$\underbrace{\tilde{\eta}_t}_{v_t} \sim N(0, \underbrace{I_m}_{Q}), \tag{B.18}$$

The time-varying component $\tilde{\beta}_t$ is initialized with mean and variance $b_0 = 0$ and $P_0 = 0.00001$. Thus, it is ensured that the time-varying parameters β_t are initialized with their starting values, collected in β_0.

The unobserved state vector $\tilde{\beta}$ is then extracted using standard forward-filtering and backward-sampling. Instead of taking the entire N x 1 observational vector y_t as the item of analysis, the approach taken here follows the univariate treatment of the multivariate series of Durbin and Koopman (2012), in which each of the elements in y_t is brought into the analysis individually. This offers significant computational gains and reduces the risk of the prediction error variance matrix becoming nonsingular during the Kalman filter procedure.

Lastly, given the components β_0, σ_η and $\tilde{\beta}$, the time-varying parameter matrix β (of dimension $T \times m$) can be constructed from (B.3).

B.1.3 Sampling the autoregressive coefficient, the unobserved component of the regression error process and the regression error variances

In this block, the autoregressive coefficient of the regression error process, ρ, the unobserved component, collected in μ, and the country-specific regression error variances, collected in σ_u^2, are drawn.

Note that, given draws of θ and β_t, ϵ_{it} and its lags are known. Thus, (B.2) breaks down to a conditional linear regression model, where ρ, μ and σ_u^2 can be obtained using a conjugate independent Normal-Inverted Gamma prior with $\rho, \mu \sim N(a_{0,\{\rho,\mu\}}, A_{0,\{\rho,\mu\}})$ and $\sigma_{u_i}^2 \sim IG(c_{0,i}, C_{0,i})$, with $c_{0,i}$ and $C_{0,i}$ being the country-specific shape and scale parameters of the prior distribution for the measurement error variance. As this prior is conjugate, it implies an (independent) Normal-Inverted Gamma posterior distribution. That is, $p(\rho, \mu|\sigma_u^2, \tilde{\beta}, \theta, y, \chi) \sim N(a_{T,\{\rho,\mu\}}, A_{T,\{\rho,\mu\}})$ and $p(\sigma_u^2|\rho, \mu, \tilde{\beta}, \theta, y, \chi) \sim IG(c_{T,i}, C_{T,i})$, $i = 1, 2, ..., N$, where $c_{T,i}$ and $C_{T,i}$ are the respective shape and scale parameters of the posterior distribution for the measurement error variance of country i. Defining ϵ as the $N \times (T-1)$ vector of stacked regressions error residuals, ϵ_{-1} as its lag, and u_i as the $(T-1) \times 1$ vector of residuals obtained from solving (B.2) for u for the respective country i, the posterior moments of the independent Normal-Inverted Gamma distribution are given by:

$$a_{T,\{\rho,\mu\}} = A_{T,\{\rho,\mu\}} \left(\chi'\Sigma^{-1}y + A_{0,\{\rho,\mu\}}^{-1} a_{0,\{\rho,\mu\}}\right) \tag{B.19}$$

$$A_{T,\{\rho,\mu\}} = \left(\chi'\Sigma^{-1}\chi + A_{0,\{\rho,\mu\}}^{-1}\right)^{-1} \tag{B.20}$$

$$c_{T,i} = c_{0,i} + (T-1)/2 \tag{B.21}$$

$$C_{T,i} = C_{0,i} + u_i'u_i/2 \tag{B.22}$$

ρ, μ and σ_u^2 can then be sampled from $p(\rho, \mu|\sigma_u^2, \tilde{\beta}, \theta, y, \chi) \sim N(a_{T,\{\rho,\mu\}}, A_{T,\{\rho,\mu\}})$ and $p(\sigma_u^2|\rho, \mu, \tilde{\beta}, \theta, y, \chi) \sim IG(c_{T,i}, C_{T,i})$ for $i = 1, 2, ..., N$.

B.2 Bayesian Vector Autoregression

Drawing heavily from Blake and Mumtaz (2015), this section lays out the BVAR with time-varying coefficients in quarterly frequency, which is used to estimate the correlations between the macroeconomic variables to draw realizations of the primary balance and public debt in the fiscal projection exercise, elaborated upon in appendix B.3.

For each of the ten sample countries, consider a time-varying coefficient VAR(p) model in reduced form, written as

$$y_t = \phi_{1,t}y_{t-1} + \phi_{2,t}y_{t-2} + ... + \phi_{p,t}y_{t-p} + u_t, \qquad u_t \sim N(0, \Sigma), \tag{B.23}$$

$$\Phi_t = \Phi_{t-1} + e_t, \qquad e_t \sim N(0, Q), \tag{B.24}$$

$t = \{1, 2, ..., T_q\}$, where T_q is the number of quarterly observations available for the VAR. y_t is a $M \times 1$ vector of demeaned endogenous variables, $\phi_{j,t}$, $j = 1, 2, ..., p$ are $M \times M$ coefficient matrices corresponding to the respective lag matrix y_{t-j} and u_t is a $M \times 1$ vector of reduced-form shocks. The time-varying parameters are collected in $\Phi_t \equiv (vec(\phi_{1,t}), vec(\phi_{2,t}), ..., vec(\phi_{p,t}))'$ and are assumed to follow random walk processes with joint error covariance matrix Q, as outlined in (B.24). The disturbances u_t and e_t are assumed to be serially uncorrelated and independent of each other for $t = 1, 2, ..., T_q$.

With Σ and each $\phi_{j,t}$, $j = 1, 2, ...p$, $t = 1, 2, ..., T_q$ being of dimension $M \times M$ and Q being of dimension $M^2p \times M^2p$, the high number of parameters to be estimated motivates Bayesian estimation techniques. Following Blake and Mumtaz (2015), a Gibbs sampling algorithm to approximate the model's joint and marginal posterior distributions is employed. The following sections briefly outline this algorithm.

B.2.1 Sampling the time-varying parameters Φ

First, the time-varying parameters, collected in Φ, are sampled from their conditional posterior distributions: Express the system of equations in (B.23) and (B.24) as

$$y_t = (I_M \otimes X_t)\Phi_t + u_t, \qquad u_t \sim N(0, \Sigma), \tag{B.25}$$

$$\Phi_t = \Phi_{t-1} + e_t, \qquad e_t \sim N(0, Q), \tag{B.26}$$

where I_M is the identity matrix of dimension M and $X_t \equiv (y'_{t-1}, y'_{t-2}, ..., y'_{t-p})$. Conditionally on the data (y), Σ, Q as well as the expected value and variance of the initial state, Φ_0, the system in equations in (B.25) and (B.26) constitutes a linear Gaussian state space model.

Following Blake and Mumtaz (2015), the expected value of Φ_0, B_0, is set to $vec(\hat{\Phi})$, where

$\hat{\Phi} = (X'X)^{-1}X'y$ is the OLS estimate of the time-invariant coefficient version of (B.23). Consistently, the variance of the initial state, $V_{B_0} = \hat{\Sigma} \otimes (X'X)^{-1}$, with $\hat{\Sigma} = \frac{(y-XB_0)'(y-XB_0)}{T-K}$, where K is the number of slope coefficients in the time-invariant VAR. Then, analogously to the respective block in the FRF algorithm, the Kalman filter can be employed to filter the unknown state Φ_t (forward-filtering step) and subsequently sample Φ_t from its conditional distribution (backward-sampling step), as described in Carter and Kohn (1994).

B.2.2 Sampling the variance-covariance matrix of the state disturbances Q

Next, the variance-covariance matrix of the state disturbances, Q, is sampled from its conditional posterior distribution. Assuming that Q follows an inverted Wishart distribution a priori and given a draw of Φ_t, Q can be sampled from an inverted Wishart distribution. That is,

$$p(Q|\Phi, \Sigma, y) \sim IW(Q_1, T_1), \tag{B.27}$$

where the posterior scale and shape parameters are given by

$$\begin{aligned} Q_1 &= (\Phi_t - \Phi_{t-1})'(\Phi_t - \Phi_{t-1}) + Q_0, \\ T_1 &= T_q + T_0. \end{aligned}$$

T_0, the prior shape parameter, is the number of observations to inform the prior. It can be interpreted as the number of fictitious observations added to the model from the prior. Q_0 is the prior scale matrix.

B.2.3 Sampling the variance-covariance matrix of the VAR disturbances Σ

In this block, the variance-covariance matrix of the VAR disturbances, Σ, is sampled from its conditional posterior distribution. In particular, conditionally on Φ_t and assuming an inverted Wishart prior for Σ, it holds that

$$p(\Sigma|\Phi, Q, y) \sim IW(\Sigma_1, T_\Sigma), \tag{B.28}$$

where the posterior scale and shape parameters are given by

$$\begin{aligned} \Sigma_1 &= u'u + \Sigma_0, \\ T_\Sigma &= T_q + T_{\Sigma_0}, \end{aligned}$$

with $u \equiv (u_1, u_2, ..., u_{T_q})$, $u_t = y_t - (I_M \otimes X_t)$, $t = 1, 2, ..., T_q$. T_{Σ_0} is the prior shape parameter, that is, the number of "artificial" observations added to the sample from the prior. Σ_0 is the prior scale matrix.

B.3 Fiscal projection algorithm

Given parameter estimates for the FRF and VAR coefficients, the algorithm used to repeatedly draw realizations - thus obtaining forecast distributions - of the primary balance- and the public debt-to-GDP ratios can be laid out. The approach presented in this section largely follows Celasun et al. (2006) and Medeiros (2012) but deviates occasionally due to the usage of Bayesian estimation techniques both in the FRF and the VAR block.

More precisely, future paths of the primary balance and the public debt ratios are repeatedly drawn from the FRF and a debt accumulation function. The primary balance forecast for country i is obtained from

$$pb_{i,T+h} = \hat{\alpha}_i + H_{i,T+h}\hat{\beta}_{T+h} + X_{i,T+h}\hat{\gamma} + \epsilon_{i,T+h}, \tag{B.29}$$

with $h = 1, 2, 3$ being the respective forecast horizon and $h = 1$ being the end-of-the-year forecast ("nowcast") of the respective vintage, $h = 2$ is the forecast for the subsequent year and $h = 3$ is the two-year-ahead forecast. $\hat{\alpha}_i$ is the estimate of the country-specific constant and can be recovered from the estimated FRF from $\hat{\alpha}_i = \bar{pb}_i - \bar{H}_i\bar{\beta} - \bar{X}_i\hat{\gamma}$, where $\bar{pb}_i$, $\bar{H}_i$ and $\bar{X}_i$ are country-specific means and $\bar{\beta} = \frac{\sum_{t=1}^{T}\hat{\beta}_t}{T}$ (barring the time-varying parameters, see for example Baltagi, 2013). The forecast for $\hat{\beta}_{T+h}$ is obtained using the non-centered parameterization and thus given by $\hat{\beta}_{T+h} = \hat{\beta}_0 + \hat{\sigma}_\eta + \hat{\tilde{\beta}}_T + \sum_{j=1}^{h}\tilde{\eta}_j$.

Note that the matrices H and X contain the fitted values of the output gap (having used an auxilliary regression to account for the variable's endogeneity as elaborated upon above) and the lagged primary balance and lagged public debt ratio. To obtain a forecast for $h = 1$, the latter two are simply their end-of-sample observations, that is, pb_{iT} and $debt_{iT}$. For the output gap on the other hand, the realization in $T+1$ is unobserved and needs to be forecasted: First, the (quarterly) $ln(GDP)$ series is forecasted using the VAR and then used to get an estimate of the cycle based on the the Hodrick-Prescott filter (where a value of $\lambda = 1600$, as conventional for quarterly data, is used). The resulting output gap in quarterly frequency is then annualized for consistency with FRF data.[19]

Note that the simulation is done R times, where R is the number of retained draws from the MCMC algorithms elaborated on in B.1 and B.2. This is convenient as for each draw $r = 1, 2, ..., R$, the respective draws of the posterior distributions - that is β_T^r (to compute β_{T+h}^r), γ^r et cetera - can be used to come up with one forecasted path of the fiscal variables. Likewise, the respective set of forecast errors ϵ_{it}^r is used to come up with

[19] To avoid the end-point problem (see e. g. Everaert and Jansen, 2017), log(output) is forecasted four quarters further into the future before computing the output gap.

the realizations of $\epsilon^r_{i,T+h}$ for each respective draw: From equation (B.2), it follows that $\epsilon^r_{i,T+h} = \mu^r_{T+h} + \rho^r \epsilon^r_{i,T+h-1} + u^r_{i,T+h}$. In the benchmark specification, μ^r_{T+h} is set to μ^r_T. However, a second alternative, of setting $\mu_{T+h} = 0$, hardly changes the results.[20] $u^r_{i,T+h}$ is obtained using bootstrapping, as in Medeiros (2012). Due to the assumption of country-specific error variances σ^2_i, $i = 1, 2, ..., N$, this is done for each country separately. Lastly, note that for $h = 1$, $\epsilon_{i,T+h-1} = \epsilon_{iT}$ is observable, such that all components to compute $\epsilon_{i,T+1}$ are known. Given $\epsilon_{i,T+1}$, $\epsilon_{i,T+2}$ can then be obtained, and so can $\epsilon_{i,T+3}$.

The public debt ratio for country i is based on the following debt accumulation equation (similar to Medeiros, 2012):

$$debt_{i,T+h} = \frac{1 + iir_{i,T+h}}{1 + (\Delta y_{i,T+h} + \pi_{i,T+h})} + pb_{i,T+h} + sfa_{i,T+h}, \tag{B.30}$$

where $debt$ is the public debt-to-GDP ratio, iir is the implicit interest rate on the debt outstanding (scaled by GDP), Δy is GDP growth, π is inflation and sfa are stock-flow adjustments of the stock of public debt (scaled by GDP), that is, one-off adjustments to the level of public debt not attributable to the other components, such as the privatization of public assets. While Δy and π forecasts can be obtained directly from the VAR, iir and sfa are taken from AMECO (see data appendix).

Note that the approach outlined here means that the real interest rate as defined above is not used in the debt simulation. Nevertheless, it is included in the VAR to adequately capture the variables' correlations. Alternative debt forecasts based on the real interest rate and not the implicit interest rate (adjusted for inflation) on average perform slightly worse than the forecasts presented here.

The AMECO database contains only point forecasts. Thus, median forecasts for each variable and horizon are computed and compared to the fixed coefficient model forecast and the EC forecast, found in the AMECO vintages.

[20] Another approach would be to forecast μ^r_{T+h}, making use of its estimates given for periods $t = 1, 2, ..., T$.

B.4 The fixed coefficient model

This section briefly outlines the fixed coefficient model (the "fixed model") that is used to judge the forecast performance of the benchmark model in the main paper. First note that the fixed model uses the same set of predictors in its FRF part and the same endogenous variables in the VAR part, as elaborated upon in section 2. The FRF is given by

$$pb_{it} = \alpha_i + X_{it}\gamma + \epsilon_{it}, \quad \epsilon_{it} \sim N(0, \sigma_\epsilon^2) \tag{B.31}$$

$i = 1, 2, ..., N, t = 1, 2, ..., T$. Note that in the fixed model, the lagged debt ratio enters the predictor matrix X, as the corresponding slope coefficient is assumed to be time-invariant. As before, X additionally contains the lagged primary balance and the output gap. Similar to Everaert and Jansen (2018), the model is estimated using a two-stage least squares instrumental variables estimator on the within-group demeaned model to account for potential endogeneity of the output gap, which is instrumented by its first and second lag.

The VAR in this case is the time-invariant coefficient pendant of equation (7):

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + u_t, \qquad u_t \sim N(0, \Sigma), \tag{B.32}$$

$t = \{1, 2, ..., T_q\}$, where T_q again is the number of quarterly observations in the VAR, y_t is a $M \times 1$ vector of demeaned endogenous variables, ϕ_j, $j = 1, 2, ..., p$ are $M \times M$ coefficient matrices corresponding to the respective lag matrix y_{t-j} and u_t is a $M \times 1$ vector of reduced-form shocks, and the model is estimated using equation-by-equation ordinary least squares.

The primary balance and debt projection block of the model mostly follows the approach outlined in section 2.4, the main difference being that, unlike for the Bayesian benchmark model, parameter uncertainty is not directly incorporated in the fiscal projection exercise.