GEORG-AUGUST-UNIVERSITY GÖTTINGEN

DOCTORAL THESIS

Designing fiscal policy

On the trade-off between macroeconomic stabilization and government debt sustainability

---

Author:
Tore DUBBERT

Supervisors:
Prof. Dr. Tino BERGER
Prof. Dr. Bernd KEMPA
Prof. Dr. Robert SCHWAGER

A thesis submitted in fulfillment of the requirements
for the doctoral degree

of the

Faculty of Business and Economics
Georg-August-University Göttingen

2023

"Personally I'm always ready to learn, although I do not always like being taught."

Winston Churchill

Executive summary

In times of reduced monetary policy leeway, COVID-19 and the energy crisis as well as long-term challenges related to demographic change and the transition towards net-zero economies, more and more policymakers turn to fiscal policy as a remedy. However, resorting to fiscal instruments as a means of fulfilling policy goals is not a free lunch, as fiscal stimulus might endanger the sustainability of public finances, thus restricting policymakers' maneuvering room (fiscal space) to address economic challenges. The fact that there is a potential downside of fiscal stimulus implies a trade-off between macroeconomic stabilization and fiscal sustainability, which should be accounted for when designing fiscal policies.

This dissertation, comprising three academic papers, deals with the trade-off along both dimensions, that is, fiscal sustainability and macroeconomic stabilization. More precisely, the first and second paper focus on the sustainability of public finances, while the third assesses the effectiveness of fiscal policies in terms of the state of the economy.

The first paper, joint with Tino Berger and Ruben Schoonackers, examines the specification of fiscal reaction functions, which capture the fiscal response (as measured by a country's primary balance) to changes in public debt and other economic conditions. We argue that the fiscal response to public debt is varying over time, finding formal evidence in favor of our time-varying parameter specification for a sample of five EU countries. We then link this non-linear fiscal response to changes in the interest rate-growth differential and the level of public debt.

In the second paper, a fiscal reaction function, featuring time-varying parameters, is embedded in a debt sustainability analysis framework to forecast the short-term evolution of the primary balance and public debt for a sample of ten OECD countries. The results suggest that debt sustainability analyses featuring time-varying parameter fiscal reaction functions perform competitively compared to a time-invariant coefficient pendant and even outperform the European Commission Economic Forecasts at certain horizons.

The third paper, joint with Tino Berger, assesses government spending policies based on their effects on the output gap in the United States. While the fiscal multiplier literature focuses on fiscal policies' effects on levels and growth rates, we can directly make statements about fiscal policy's impact on the business cycle by means of the Beveridge-Nelson decomposition. We find that the dosage of expansionary fiscal policy is key: Pronounced fiscal stimulus does increase the output gap, but can lead to an overheating economy and increasing public debt levels. While this finding is in line with standard macroeconomic models, our approach can guide policymakers in quantifying a policy's impact on the business cycle.

Kurzfassung

In Zeiten mangelnden geldpolitischen Spielraums, COVID-19 und der Energiekrise sowie langfristigen Herausforderungen im Zuge des demographischen Wandels und dem Übergang hin zu einer CO2-neutralen Wirtschaft gewinnt die Fiskalpolitik zunehmend an wirtschaftspolitischer Bedeutung. Fiskalische Maßnahmen sind allerdings mit Kosten verbunden: So können Konjunkturprogramme die Tragfähigkeit öffentlicher Finanzen gefährden und damit den fiskalpolitischen Spielraum, wirtschaftspolitische Ziele zu verfolgen, einschränken. Diese Kehrseite von Konjunkturprogrammen impliziert einen möglichen Zielkonflikt zwischen makroökonomischer Stabilisierung und fiskalischer Tragfähigkeit, für den Entscheidungsträger bei der Strukturierung von Fiskalpaketen Sorge tragen sollten.

Diese Dissertation besteht aus drei Forschungspapieren, die sich diesem Zielkonflikt sowohl im Hinblick auf die fiskalische Tragfähigkeit (Papiere eins und zwei) als auch auf die makroökonomische Stabilisierung (Papier drei) widmen.

Das erste Papier (mit Tino Berger und Ruben Schoonackers) untersucht die Spezifikation von Fiskalreaktionsfunktionen, die die fiskalische Antwort (Veränderung des Primärsaldos eines Landes) auf eine Veränderung der Staatsverschuldung und anderer ökonomischer Größen messen. Wir argumentieren, dass die fiskalische Reaktion auf eine Veränderung der Staatsverschuldung zeitvariabel ist, und finden für unsere Stichprobe von fünf europäischen Ländern formale empirische Evidenz für diese Zeitvariabilität. In einem zweiten Schritt finden wir, dass diese nicht-lineare fiskalische Reaktion durch Änderungen des Zins-Wachstums-Differentials sowie die Höhe der Staatsverschuldung getrieben werden.

Im zweiten Papier werden Fiskalreaktionsfunktionen mit zeitvariablen Parametern im Rahmen der Schuldentragfähigkeitsanalyse eingesetzt, um den Primärsaldo und die Staatsverschuldung in der kurzen Frist für zehn OECD-Länder zu prognostizieren. Schuldentragfähigkeitsanalysen dieser Art liefern „wettbewerbsfähige“ Prognosen, insbesondere im Vergleich mit entsprechenden Modellen ohne zeitvariable Parameter. Auch im Vergleich zu den „Economic Forecasts“ der Europäischen Kommission liegen für einige Prognosehorizonte Performance-Vorteile vor.

Im dritten Papier (mit Tino Berger) werden die Effekte von Staatsausgabenpolitik auf die Outputlücke in den USA untersucht. Über die Nutzung der Beveridge-Nelson-Zerlegung treffen wir direkte Aussagen über fiskalpolitische Effekte auf den Konjunkturzyklus, während der Fokus der Literatur fiskalischer Multiplikatoren auf Level- und Wachstumsvariablen liegt. Unseren Ergebnissen zufolge ist die Dosierung von Konjunkturpaketen entscheidend: Expansive Fiskalpolitik in Krisen führt zu einer schnelleren Schließung der Outputlücke, kann aber zu Überhitzung und hoher Staatsverschuldung führen. Während dieser Befund den Ergebnissen von Standardmodellen der Makroökonomik entspricht, gibt unser Ansatz Entscheidungsträgern ein Modell an die Hand, das die Quantifizierung der Effekte einer fiskalischen Maßnahme auf den Konjunkturzyklus ermöglicht.

Acknowledgements

Getting to this point has been quite the journey. Many people have accompanied me on the way towards the finishing line of this dissertation, and I want to dearly thank all of them. That said, I want to thank the following people in particular:

First, I want to express my deep gratitude towards my first supervisor, Tino Berger, for having been a patient mentor. Our countless discussions and his support have been crucial driving forces for succeeding with this dissertation. I have learned a lot on how to tackle immense projects such as this, something that has become invaluable to me not only in a professional but also in a personal sense.

Moreover, I am truly grateful for the support of my second supervisor, Bernd Kempa, not only for having made me a part of the great team at the Institute of International Economics at the University of Münster, but for deepening my understanding of macroeconomics and always being a helpful listener to various project ideas (and problems).

I am also very thankful for the help of my third supervisor, Robert Schwager, for having joined my committee, for his helpful comments especially on chapter 2 of this dissertation, and for having been one of my first economics professors during my Bachelor studies at the University of Göttingen, inspiring me early on to pursue topics in the subject area of public finance.

I am grateful to my co-author Ruben Schoonackers, who has greatly improved the work of the second chapter of this dissertation and, by implication, the third. I have highly profited from our professional exchange.

I want to thank all my colleagues and former colleagues, starting with Julia Richter, for her great support and encouragement from the very beginning. Thanks to Markus Ahlborn for the time we shared in Göttingen, Carina Börger and Florian Hüpper for the years we spent together at the University of Münster – I have truly enjoyed them. The same goes for Feina Zou and Adrian Schröder, who have joined the team more recently, and for Nazmus Sadat Khan, for our time together in Münster and our collaboration on fiscal policy questions at The World Bank. Thanks to Philipp Hansen for great discussions and support. I want to express my gratitude towards all other (former) colleagues and friends, it would take another dissertation to list them all.

Finally, I want to thank my family, who has (mostly) succeeded in the nearly impossible task of keeping me sane, with unlimited support. I want to thank my parents, my sister and my wife, who have made this possible.

Contents

List of Figures

List of Tables

To my family.

Chapter 1

Introduction

After the Great Recession hit the United States in late 2007, policymakers engaged in massive monetary stimulus. In fact, so massive was the stimulus that the Federal Funds Rate swiftly approached the zero-lower bound – notwithstanding levels above 5% in mid-2007 – where it would remain for many years. As a result, the Fed's leeway to stimulate the economy by conventional means became severely limited. Due to similar developments in many industrialized economies, countries increasingly turned to fiscal policy as a means of macroeconomic stabilization. In the words of former Fed chairman Ben Bernanke: *"Monetary policy has less room to maneuver when interest rates are close to zero, while expansionary fiscal policy is likely both more effective and less costly in terms of increased debt burden when interest rates are pinned at low levels."*[1]

Clearly though, resorting to fiscal measures for stabilization purposes is far from being a free lunch. In the above words of Bernanke: Expansionary fiscal policy being *less* costly near the zero-lower bound (in terms of the resulting debt burden) does not mean it is *without* costs. More precisely, fiscal policy's leeway to fulfill macroeconomic stabilization goals is constrained by the magnitude of the public debt position and the creditors' willingness to finance it: If financial markets question the sustainability of a country's debt and deficit positions, higher refinancing costs might arise, further weighing on the sustainability of public finances and starting a vicious cycle.

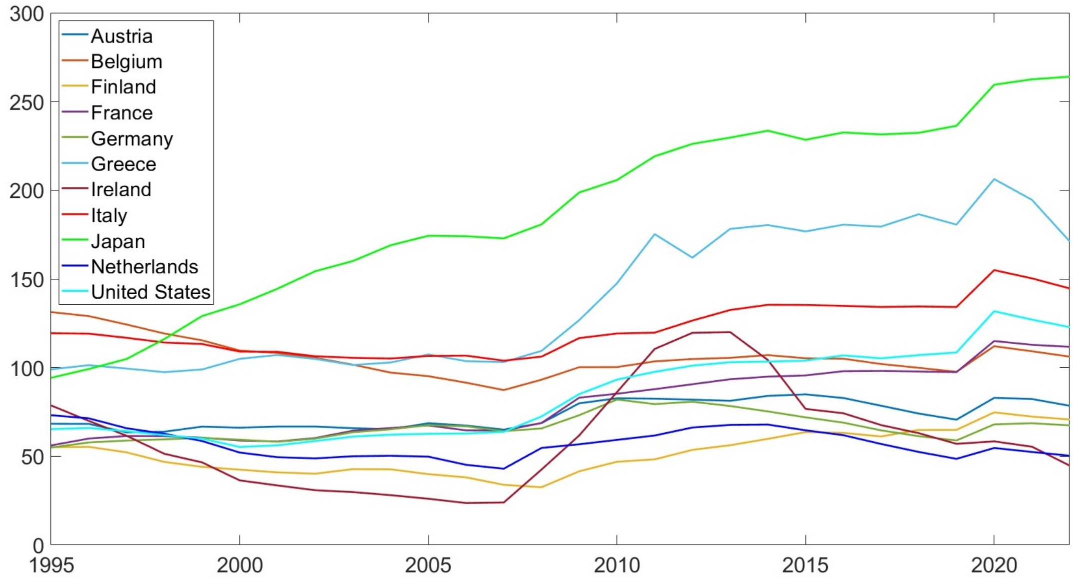
Figure 1.1 depicts the evolution of public debt-to-GDP ratios ("debt ratios") for a selection of industrialized economies from 1995 to 2022.[2] First, note that debt ratios have been rising steadily. In eight out of eleven countries, debt ratios have increased from 1995 to 2022, in some cases drastically: For France, Greece, Japan and the United States, this increase is well above 50 percentage points. Rising debt ratios – both in the past and in the future – are not surprising, given long-term challenges in many industrialized economies such as an increasing financial burden related to demographic change as well as the (at least initially) costly transition towards low-carbon economies.

Second, attempts to stabilize the economy in response to major economic crises seem to have caused a deterioration in fiscal positions: Debt ratios in all displayed

[1] https://www.federalreserve.gov/newsevents/speech/bernanke20140103a.htm.

[2] The selection of countries coincides with the sample countries chosen in the following chapters of this dissertation.

FIGURE 1.1: Evolution of public debt ratios in industrialized economies



Notes: This figure shows the evolution of public debt in percent of GDP for a selection of industrialized economies from 1995 to 2022. The countries displayed are those countries used in the samples of chapters 2, 3 and 4. Data source: AMECO database.

countries have risen in the aftermath of the Global Financial Crisis – and in some cases strongly – thus reducing governments' "fiscal space".[3] Similar dynamics can be seen for the COVID-19 crisis, where debt ratios have, at least in 2020, increased in all depicted countries. Thus – while effective at the zero-lower bound according to Bernanke – excessive fiscal stimulus might induce high costs related to a growing public debt stock, at least partially offsetting the benefit from the fiscal expansion.[4]

The fact that there is a potential downside from fiscal stimulus implies a *trade-off between macroeconomic stabilization and fiscal sustainability goals* (see e. g. Hauptmeier and Kamps, 2022). Any prudent fiscal authority should properly take fiscal (public debt) sustainability into account when designing a stimulus package. This dissertation attempts to contribute with respect to these two competing fiscal policy goals: Chapters 2 and 3 address the adequate assessment of public debt sustainability. Chapter 4 analyzes the efficacy of fiscal policy in terms of macroeconomic stabilization. These chapters may be seen as steps towards a comprehensive, model-based analysis of an "optimal" fiscal policy design.[5]

[3]See Ghosh et al. (2013), who state that fiscal space can be understood as the distance of the actual public debt level from a limit beyond which the debt position becomes unsustainable.

[4]Note that increasing debt ratios can also come about due to falling GDP levels. However, due to the major fiscal stimulus packages implemented in the depicted countries in the aftermath of the Global Financial Crisis as well as in response to the COVID-19 shock, fiscal stimulus is certainly at least partly to blame for the increase.

[5]It should be noted, though, that whether fiscal stimulus indeed worsens a country's debt position is not necessarily straightforward, as growth effects from fiscal stimulus might (partially or fully) offset the deficit-induced debt ratio hike (see e. g. McCausland and Theodossiou, 2015). Nevertheless, running fiscal deficits can severely jeopardize fiscal sustainability by reducing creditors' trust in the

Chapter 2, which is joint work with Tino Berger (University of Göttingen) and Ruben Schoonackers (National Bank of Belgium, Vrije Universiteit Brussel), attempts to contribute along the lines of public debt sustainability by addressing adequate model selection. Fiscal empiricists have frequently used *fiscal reaction functions* (FRFs) – behavioral equations for fiscal authorities – to assess public debt sustainability (see especially the pioneering work of Bohn, 1995, 1998, 2008). Bohn argues that, if a government on average increases its primary balance (its policy variable) in response to rising government debt, this would be a sufficient condition for public debt to be sustainable. As outlined in Everaert and Jansen (2018), FRFs can conveniently be embedded in *stochastic debt sustainability analysis* (SDSA) frameworks, where future primary balance and public debt paths are simulated to judge the sustainability of public finances by means of *fan charts*, depicting projected distributions of the variables of interest, allowing the researcher or policymaker to make probabilistic statements about the variables' future evolution (e. g. probabilities for certain thresholds being reached; see e. g. Medeiros, 2012, Celasun et al., 2006). Everaert and Jansen (2018) stress that, since FRFs play an important role in (stochastic) debt sustainability analysis, it is crucial they be appropriately specified. Only if this is the case will it be possible to adequately assess fiscal sustainability.

Addressing this argument on the importance of correctly specifying FRFs, we argue that there is reason to believe that the fiscal reaction to public debt (by means of the primary balance) may vary over time. First, such time variation might come about due to differences in fiscal responsiveness at varying levels of debt. As argued by Ghosh et al. (2013), the fiscal reaction to debt might be increasing with the level of debt up to a certain point, from which onward the primary balance cannot keep up with ever-rising amounts of public debt – "fiscal fatigue" settles in. Second, debt dynamics crucially depend on the interest rate-growth differential: While higher growth implies higher tax generation capacities and lower unemployment benefit expenses (as well as a higher denominator in the debt ratio), higher interest rates on the debt, demanded by financial markets, imply higher debt service costs (see e. g. Blanchard, 2019). Thus, both variations in growth and interest rates render any given level of debt more or less sustainable, potentially implying the necessity to adjust fiscal behavior. Our baseline model thus features a time-varying primary balance reaction to public debt. Additionally, we use a formal model selection algorithm to assess the "probability" of our specification, given data and prior. Thus, we directly address the above-mentioned importance of correctly specifying FRFs, given their role in debt sustainability analysis. We find strong evidence for a time-varying fiscal reaction to public debt. In a second step, we seek to explain this observed time variation, enriching our model by including the interest rate-growth differential as well as the level of public debt itself as explanatory variables for the time-varying fiscal reaction to debt. We find that both covariates have some explanatory power,

sustainability of a rising public debt stock, thus increasing debt service costs and weighing on growth and debt sustainability.

while other (not considered) determinants appear to play a role, too.

Just as chapter 2, chapter 3 (a revised version is accepted for publication in *Applied Economics*) is concerned with model specification, thus attempting to contribute along the lines of proper fiscal sustainability evaluation. Unlike in chapter 2, where the focus is on the correct FRF specification based on models' in-sample fit, in chapter 3 I am focusing on the out-of-sample forecasting performance of SDSA frameworks that feature time-varying parameter FRFs, using vintage datasets. Such "real-time" fiscal forecasting exercises to judge debt sustainability analysis models have not been treated very prominently in the relevant literature, certainly at least partly due to difficulties in retrieving relevant data. This chapter tries to fill this gap. Various model specifications are judged based on whether they have produced primary balance and public debt forecasts close to actual ex-post values. For this purpose, SDSA models in spirit of Medeiros (2012) and Celasun et al. (2006) are employed. That is, next to a FRF, a vector autoregressive model (VAR) to capture the correlations between macroeconomic variables that affect the fiscal variables is included in the framework. Unlike in Medeiros (2012) and Celasun et al. (2006), FRFs and VARs in the specifications in chapter 3 feature time-varying parameters, thus allowing for changes in underlying relationships over time. Assume, for example, that the actual data-generating process of the fiscal reaction to public debt in the FRF is time-varying. If, in the current period, the true parameter is significantly below its mean and will stay there for some time, using a fixed coefficient FRF within a debt sustainability framework can lead to sustainability risks being severely underestimated. Using SDSA frameworks with time-varying parameters can mitigate misleading inference resulting from such situations.

I find that SDSA models featuring time-varying parameter FRFs and VARs perform competitively in terms of mean squared error and forecast bias against a state-of-the-art (Medeiros, 2012-style) SDSA model without time-varying parameters and the European Commission *Economic Forecasts* at short horizons (up to two years ahead), in particular for public debt. Thus, SDSA frameworks featuring FRFs (and VARs) with time-varying parameters should serve as a complementary tool to debt sustainability analyses conducted in academia and at policy institutions.

While chapters 2 and 3 address the debt sustainability goal of fiscal policy, chapter 4, which is joint work with Tino Berger, looks at the economic stabilization part of the trade-off: Given that (massive) fiscal stimulus potentially weighs on fiscal sustainability, how effective were certain stimulus programs in reaching policy goals? To answer this question, we look into the efficacy of hypothetical government spending scenarios in terms of stabilizing the US economy in major crises in history. Fiscal policy effects on the macroeconomy are usually quantified by means of fiscal multipliers. For example, an answer to the question "how much Dollar in GDP are generated from a one Dollar government spending expansion?" is provided (see e. g. Caldara and Kamps, 2017 for a detailed overview on fiscal multipliers). However, the literature on fiscal multipliers is vague when it comes to quantifying effects on the business cycle. Looking into business cycle effects can be fruitful since, next to adversely affecting public finances, a fiscal expansion might lead to an overheating of the economy with distortionary effects arising from the implied increase in inflation. Thus, assessing effects on the business cycle provides a clearer picture on the state of the economy compared to fiscal multipliers and is therefore the more policy-relevant measure.

We contribute to the debate on fiscal policy efficacy in terms of business cycle stabilization on the government spending side. We identify the business cycle from a multivariate Beveridge-Nelson decomposition based on a VAR in spirit of Morley and Wong (2020) and look into counterfactual scenarios for government spending, asking "what would the path of the business cycle have looked like if government spending growth had been x% higher for some periods?". We look into four major economic crises in US history, finding that the "dosage" of the stimulus packages is crucial. More precisely, higher government spending growth does have a positive effect on the business cycle. That said, there is a real danger from "overspending", leading to pronounced overheating and debt sustainability risks, the latter of which reduces the fiscal space for tackling future crises. Thus, while government spending increases can certainly be used to overcome adverse macroeconomic conditions, these downside risks should be taken into account when designing a fiscal stimulus package. While these findings are well-known, we provide policymakers with a framework to quantify opportunities and challenges of a planned stimulus program within a model that does not require restrictive identification assumptions to identify fiscal shocks.

As outlined above, this dissertation contributes along the lines of the two potentially competing fiscal policy goals of macroeconomic stabilization and public debt sustainability. Chapters 2 and 3 focus on the debt stabilization goal, contributing to the literature by operationalizing in- and out-of-sample model selection in a fiscal policy context and providing formal evidence in favor of debt sustainability frameworks featuring time-varying parameter models. Chapter 4 analyzes the effectiveness of fiscal policy in terms of business cycle stabilization, providing a tool to quantify fiscal policy effects on the business cycle that is independent of the identification strategy of fiscal shocks. Future work could combine the components of these chapters in a holistic approach to designing effective, sustainable fiscal policy.

Bibliography

Blanchard, O. (2019). Public debt and low interest rates. *American Economic Review*, 109(4):1197–1229.

Bohn, H. (1995). The sustainability of budget deficits in a stochastic economy. *Journal of Money, Credit and Banking*, 27(1):257–271.

Bohn, H. (1998). The behavior of U. S. public debt and deficits. *The Quarterly Journal of Economics*, 113(3):949–963.

Bohn, H. (2008). The sustainability of fiscal policy in the united states. *Sustainability of public debt*, pages 15–49.

Caldara, D. and Kamps, C. (2017). The analytics of SVARs: A unified framework to measure fiscal multipliers. *The Review of Economic Studies*, 84(3):1015–1040.

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.

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

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.

Hauptmeier, S. and Kamps, C. (2022). Debt policies in the aftermath of covid-19 — the sgp's debt benchmark revisited. *European Journal of Political Economy*, 75:102187.

McCausland, W. D. and Theodossiou, I. (2015). The consequences of fiscal stimulus on public debt: a historical perspective. *Cambridge Journal of Economics*, 40(4):1103–1116.

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.

Morley, J. and Wong, B. (2020). Estimating and accounting for the output gap with large bayesian vector autoregressions. *Journal of Applied Econometrics*, 35(1):1–18.

Chapter 2

Fiscal prudence: It's all in the timing – Estimating time-varying fiscal policy reaction functions for core EU countries

with Tino BERGER[1] and Ruben SCHOONACKERS[2]

Abstract

When estimating fiscal policy reaction functions (FRF), the literature has well recognized the importance of non-linearities. However, there is yet very little attempt to formally test for the presence and potential sources of a non-linear fiscal responsiveness. In this paper we address this gap by formally addressing model specification of the FRF in a panel of five EU countries. Employing a Bayesian stochastic model specification search algorithm, we provide formal evidence for time-varying fiscal prudence over the last 50 years. The primary balance responsiveness exhibits smooth but significant variation over time and thus confirms the necessity of a non-linear model. Moreover, the extended results show that dynamics can be partially linked to the interest rate growth differential and the level of public debt itself. However, no clear evidence is found in favor of the fiscal fatigue proposition.

Keywords: Fiscal reaction function, time-varying parameters, state-space models, MCMC, stochastic model specification search

JEL Codes: E62, H63, C11, C52

[1]University of Göttingen.
[2]National Bank of Belgium, Vrije Universiteit Brussel.

2.1 Introduction

With the policy rate at the zero lower bound the ability of monetary policy to encourage economic growth is limited. As a result, the role of fiscal policy to stabilize the economy has become increasingly important in recent years. Moreover, for euro area countries fiscal policy is the only instrument that allows for country-specific economic policy. There is an extensive and ongoing discussion in the literature on fiscal sustainability. While the importance of safe debt levels is well recognized, recent events seriously challenge the sustainability of public finances.

To counteract the consequences of the COVID-19 induced recession, governments responded with unprecedented spending. In the euro area, the government deficit increased to 8.8% of GDP in 2020 and the gross public debt ratio reached a level of 101.7% of GDP. With low economic growth and prolonged stimulus packages, these numbers are likely not to decline quickly over the next few years. Additionally, rising age-related public expenditures and low expected potential growth further challenge the sustainability of public finances.

In assessing sovereign vulnerabilities, stochastic debt sustainability analysis (DSA) frameworks play an important role. They make use of simulated stochastic debt trajectories that reflect the interplay of model-based projections for relevant macroeconomic variables with an expected fiscal policy response, based on the estimation of a fiscal policy reaction function (FRF) that describes how the primary balance responds to changes in public debt. Besides its importance for DSA, the estimation of a FRF yields information on the type and strength of fiscal policy reactions governments had in the past and can be helpful in providing signals for potential future sustainability issues. In order to obtain reliable debt projections, a correctly specified FRF is thus essential.

The FRF literature has well recognized the importance of non-linearities for correctly specifying the FRF. Ghosh et al. (2013), for example, find strong evidence for the existence of a non-linear FRF that exhibits fiscal fatigue which is very well approximated by a cubic relationship between public debt and the primary balance. When the level of inherited debt increases, the primary balance responsiveness also increases but eventually starts to decrease and at high levels of debt finally becomes negative. Recent studies such as Fournier and Fall (2017) attempt to derive these thresholds endogenously by employing regime-switching models, while others such as Weichenrieder and Zimmer (2014), explicitly link the fiscal reaction to a specific event such as Euro membership.

Despite the emerging consensus regarding the importance of correctly modeling the FRF, there is yet very little attempt to formally test for the presence and the potential source of non-linearities. In this paper we address this gap in the literature by formally addressing model specification of the FRF in a panel of five EU countries. Specifically, we employ a Bayesian model specification search to test for time-variation in the responsiveness in the primary balance to the gross public debt ratio. The responsiveness parameter is allowed to vary according to a random walk and thus allows for various forms of non-linear reaction. We then test how much of the time-variation can be explained by the interest rate growth differential and the lagged squared debt ratio. The latter essentially tests for the fiscal fatigue proposition of Ghosh et al. (2013).

We find strong evidence for time-variation in the FRF over the last 50 years. The primary balance responsiveness to debt exhibits smooth but significant variation over time and thus confirms the necessity of a non-linear model. The dynamics can be partially linked to the interest rate growth differential. Less evidence is found in favor of the fiscal fatigue proposition.

The remainder of the paper is structured as follows. In Section 2.2, we discuss the empirical setup. More specifically, we focus on the role of FRFs in debt sustainability analysis and we elaborate on our empirical specification. Section 2.3 focuses on the econometric approach for estimating our FRF, while in Section 2.4 results of our empirical analysis are discussed. Section 2.5 concludes.

2.2 Empirical setup

2.2.1 Estimating a FRF: The basics

Is government policy in line with fiscal solvency? This is a question that features promintently in the academic and policy debate and boils down to assessing whether the debt-to-GDP ratio belongs to a dynamically stable trajectory.[3]

To analyze fiscal solvency, recall the public debt accumulation equation,

$$\Delta d_t \equiv d_t - d_{t-1} = \frac{r_t - g_t}{1 + g_t} d_{t-1} - pb_t, \tag{2.1}$$

where d_t and pb_t respectively stand for the debt-to-GDP and the primary balance ratio in period t. The interest rate on the outstanding amount of debt is represented by r_t whereas nominal GDP growth equals g_t. From equation (2.1) one can immediately see that debt dynamics are driven by two opposing forces, (i) the interest-rate growth differential (IRGD) $(r_t - g_t)$ and (ii) the primary balance.

If then the fiscal reaction to debt is represented by

$$pb_t = \beta d_{t-1}, \tag{2.2}$$

one can derive that to ensure a dynamically stable public debt trajectory, i. e. a mean-reverting public debt ratio, on average the following condition needs to hold:

$$\beta > \frac{r - g}{1 + g} \tag{2.3}$$

[3] A very interesting overview of key economic principles and statistical methods used in debt sustainability analysis is given in Debrun et al. (2019). This section is indebted to their lecture.

The fiscal reaction to debt – and estimating a FRF – can thus be used to assess the sustainability of public finances.

In a seminal paper, Bohn (1998) was the first to analyze this kind of FRF. More specifically, Bohn's (1998) model-based sustainability test (MBS) consists of estimating

$$pb_t = \beta d_{t-1} + X_t \gamma + \epsilon_t, \tag{2.4}$$

where X captures a set of other determinants explaining the evolution in the primary balance and ϵ_t represents a white noise error term. Bohn (1998) showed that, under a set of regularity conditions, a positive primary balance reaction to changes in the debt ratio (i. e. $\beta > 0$) is sufficient evidence for an economy to be fiscally sustainable and satisfying its intertemporal budget constraint. Applications of Bohn's MBS differ mainly regarding the covariates included in X and the empirical setting, i. e. the country and time coverage. Two interesting and comprehensive overviews of existing FRF studies are given by Berti et al. (2016) and Checherita-Westphal and Žd'árek (2017).

Another major contribution to the FRF literature is the often cited paper of Ghosh et al. (2013). In their analysis, the authors argue that an average positive fiscal reaction to debt ($\beta > 0$) should be labeled as a "weak" sustainability criterion as this implies that an ever-increasing debt-to-GDP ratio is not excluded.[4] For instance, this is the case if the increase in the primary balance is lower than the IRGD. They advocate a stricter sustainability criterion – the public debt ratio converging to some finite proportion of GDP – and argue that a sufficient condition for this is a primary balance reaction that on average exceeds the interest-rate growth differential.

A key issue in the fiscal austerity debate – and in the estimation of FRFs – is whether the degree of fiscal responsiveness to public debt changes with the level of debt. Specifically, the hypothesis of fiscal fatigue has been tested. Ghosh et al. (2013), for example, find strong support for a non-linear relationship between the primary balance and the lagged debt ratio that exhibits fiscal fatigue. More precisely, they find a cubic relation: At low levels of debt, the relationship between the primary balance and debt is barely existent. But as debt increases, the primary balance reacts positively and increases (more than proportionally) with the stock of debt. Eventually, the response starts to weaken and even decreases at very high levels of debt. Thus, at very high debt levels, the fiscal effort – in the form of raising extra taxes or cutting primary spending – required to "keep up" with debt becomes unfeasable and/or undesirable. In more recent publications, Fournier and Fall (2017) confirm the fiscal fatigue property for a group of OECD countries, while Everaert and Jansen

[4]This is in contrast to Bohn (1998), who considers $\beta > 0$ – under reasonable regularity conditions – to be a sufficient condition to meet fiscal solvency.

(2018) find it not to be a general characteristic of the FRF when allowing for country-specific fiscal reactions to public debt.[5] Testing for the presence of fiscal fatigue is important as this implies the existence of a debt level – the debt limit – where the debt dynamics become explosive and the government will inevitably default.[6]

By analyzing the fiscal fatigue property, the literature obviously recognizes that the primary balance reaction could display significant variation over time. However, accounting for a non-linear relationship between the primary balance and public debt by adding potencies of the debt variable to the regression equation – as in Ghosh et al. (2013), where squared and cubic debt terms are added – only allows for a very specific, deterministic source of time variation. A time-varying policy response could also be due to different sources, such as the response to a changing IRGD. Others, like Weichenrieder and Zimmer (2014) have tried to link fiscal responsiveness to Euro membership.[7] More generally, explicitly allowing and testing for significant time variation in β and – in a stochastic approach – linking it to a set of potential determinants could be very relevant. In the remainder of the paper, we will analyze this type of FRF.

The importance of analyzing whether or not there is significant time variation in β is also recognized by Debrun et al. (2019) and relates to the long-term perspective when using FRFs as a test of fiscal sustainability. Debrun et al. (2019) state that in order for the outcome of FRF-based sustainability tests to be meaningful, the fiscal policy response to lagged public debt must be sufficiently systematic and stable over time. In other words, if the response is positive for a couple of years but becomes insignificant afterwards, no clear-cut indication can be given in terms of whether fiscal policy is sustainable or not – unless the time variation can be linked to conditions that are in correspondence with fiscal solvency.

So far, only few studies have modeled time-varying FRFs in a stochastic way. Among the notable exceptions are Legrenzi and Milas (2013), who employ a regime-switching model to investigate the relationship of the primary balance with debt and other variables for Greece, Ireland, Portugal and Spain. More closely related to our approach is Burger et al. (2011), who cast their FRF, featuring a time-varying fiscal response to debt, in state-space form, finding a time-dependent fiscal reaction for their South African sample. However, they do not formally test for the presence of time variation.

[5]As Everaert and Jansen (2018) note, this it at least not the case for the range of debt levels observed in their sample of 21 OECD countries over the period 1970-2014.

[6]This is the case even when a risk-free interest rate is assumed and thus abstraction is made from the endogeneity of the risk premium on government debt.

[7]In their panel regression, Weichenrieder and Zimmer (2014) find a systematic reduction in fiscal prudence when becoming a Eurozone member. However, their result is not robust to excluding Greece from the sample, which leads to the conclusion that Eurozone membership does not significantly decrease fiscal prudence.

2.2.2 The empirical specification

To identify a governments' fiscal reaction to a changing debt ratio, we employ a dynamic specification of the FRF. As noted by, amongst others, Everaert and Jansen (2018), the highly politicized nature of the public budgeting proces makes it hard to react immediately to changes in debt and other economic conditions. Moreover, as the implementation of budgetary policies and new fiscal measures takes time, the primary balance pb_{it} is considered to be a very persistent series. A dynamic specification is thus highly justified,[8]

$$pb_{it} = \alpha_i + \delta_t + \phi pb_{i,t-1} + \beta_t d_{i,t-1} + X_{it}\gamma + \epsilon_{it}, \qquad \epsilon_{it} \sim N(0, \sigma_\epsilon^2), \tag{2.5}$$

where subscripts i and t respectively denote the ith country and tth period.

Country fixed effects α_i are included to account for country-specific time-invariant factors that affect the primary balance and are not included in X_{it}. Next to that, time-fixed effects δ_t are also present to control for the impact of global economic shocks such as the Financial and Economic Crisis of 2007-2008.

The vector X_{it} represents a $1xk$ vector of k explanatory variables that can have a direct impact on pb_{it}. The set of variables present in X_{it} resembles standard choices in the literature (see, among others, Ghosh et al., 2013, Everaert and Jansen, 2018, Berti et al., 2016, Checherita-Westphal and Žd'árek, 2017). A first variable included in X_{it} is a measure of the economic cycle, the output gap (OG_{it}), to control for the reaction of fiscal variables to the business cyle. Next to that, a measure of inflation (π_{it}) is added to account for bracket creep effects, i. e. in a progressive tax system where tax brackets are not fully indexed, rising inflation induces more than proportional changes in tax revenue (see, amongst others, Saez, 2003). An election cycle dummy variable ($elec_{it}$) is also taken into account to control for the possible presence of a political budget cycle, meaning that governments tend to increase their spending in election years to increase the probability of being re-elected (see for example, Debrun et al., 2008). Finally, the implicit interest rate on the outstanding amount of public debt (r_{it}) is included to capture potential offsetting changes in the primary balance due to changing debt services in order to reach a nominal balance target.

Naturally, we are mainly interested in the relation between pb_{it} and $d_{i,t-1}$ which represents the one-period-lagged debt-to-GDP ratio. We consider this relationship to be time-varying and allow the parameter β_t to change over time according to a random walk process,

$$\beta_t = \beta_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2). \tag{2.6}$$

[8]By allowing for a dynamic specification, potential spurious regression issues are circumvented, i .e. by adding the lagged dependent variable a potential random walk of the dependent variable is nested in the model, leading to estimates that are valid even in the case of non-stationarity.

This specification allows for a very flexible evolution of the parameter β_t over time. A random walk process is particularly convenient to capture smooth transition and structural changes. As such, by letting β_t evolve according to a random walk, we allow for frequent changes of a government's fiscal reaction to the debt ratio without forcing parameters to change.[9]

Consequently, the model that will be estimated and tested for the presence of time variation in the reaction of the primary balance to changes in the lagged public debt ratio is represented by equations (2.5)-(2.6).

Conditional on finding evidence for time variation in β_t, a model extension will be considered where we account for the possibility of observed variables playing a role in the determination of the time-varying path of β_t. More specifically, the following extended model will be estimated:

$$pb_{it} = \alpha_i + \delta_t + \phi pb_{i,t-1} + \beta_{it} d_{i,t-1} + X_{it}\gamma + \epsilon_{it}, \qquad \epsilon_{it} \sim N(0, \sigma_\epsilon^2). \tag{2.7}$$

$$\beta_{it} = \beta_t^* + G_{it}\kappa \tag{2.8}$$

$$\beta_t^* = \beta_{t-1}^* + \eta_t^*, \qquad \eta_t^* \sim N(0, \sigma_{\eta^*}^2). \tag{2.9}$$

In the model (2.7)-(2.9), the primary balance responsiveness to public debt (β_{it}) is modeled as a linear combination of a random walk component β_t^* and a set of explanatory variables G_{it}. A first variable to be included in G_{it} is the lagged public debt ratio ($d_{i,t-1}$) and its square term. By doing so, we explicitly take into account the possible presence of non-linearities in the relation between public debt and the primary balance. Moreover, by including the squared debt term, which in fact implies a cubic relation between pb_{it} and $d_{i,t-1}$, we are able to test the fiscal fatigue property as found in the seminal paper of Ghosh et al. (2013). An advantage of our approach, compared to Ghosh et al. (2013) and others, is that we do not consider the relationship to be deterministic but account for it in a stochastic way, i. e. acknowledging that other factors also drive the relation between debt and the primary balance. As such, if, empirically, a non-linear relation is found, this is not the result of ignoring other potential sources of time variation.

As can be seen from equation (2.1), the IRGD plays an essential role in public debt management. A declining, but positive, IRGD differential reduces the primary surplus needed for debt stabilization. When the IRGD becomes negative, the debt ratio can decline even when running a primary deficit. Consequently, in this type of situation one could argue that the government's fiscal responsiveness to public debt fades. Therefore, the IRGD will be included in G_{it}. More specifically, we use the IRGD at the beginning of the period, i. e. $IRGD_{i,t-1}$, as this is the relevant indicator that impacts on the discretionary fiscal policy behavior in period t.

[9] The fixed parameter specification constitutes a special case of the random walk process, with the time-varying parameter β_t in equation (2.6) being constant for $\sigma_\eta^2 = 0$.

Note that other possible determinants of the time-varying responsiveness of the primary balance to public debt, not present in G_{it}, are captured by the random walk process β_t^*. Testing for the presence of significant time variation in β_t^* can therefore be very interesting. If no significant time variation is found, this implies the variables in G_{it} fully explain the existing time variation in β_t, $\sigma^2_{\eta^*}$ will equal zero and β_t^* becomes a constant.

2.2.3 Data and Sources

In the empirical analysis a balanced panel of yearly data is used for 5 core EMU countries, covering a period from 1970-2019. The included countries are Austria, Belgium, France, Germany and the Netherlands. The choice of countries is limited to – as we believe – a somewhat homogeneous group of countries. The reason being that as our model (2.5)-(2.6) indicates, we assume a homogenous primary balance reaction to the included covariates (see section 2.3 for details on why this is important). Extending the number of countries would probably contradict that assumption.

The main data source for our analysis is the AMECO database. For the fiscal variables pb_{it}, $d_{i,t-1}$ and r_{it}, data before 1995 are retropolated using the historical public finance database prepared by Mauro et al. (2015). The inflation variable π_{it} is constructed using the GDP deflator while the $IRGD$ represents the difference between r_{it} and the nominal GDP growth rate. The OG_{it} is also taken from AMECO and is calculated using the commonly agreed production function methodology (for more details on this methodology, see Havik et al., 2014). Finally, data on $elec_{it}$ are taken from the Database of Political Institutions (DPI) version 2017, and where necessary complemented with data from older versions.

2.3 Estimation methodology

To estimate our empirical model represented by equations (2.5) - (2.6) or the extended model, (2.7) - (2.9), a number of methodological or econometrical choices need to be made. In what follows, some details are provided on these choices and on the actual methodology used to estimate the different models.

2.3.1 The choice of a homogeneous panel specification

When considering debt-to-GDP ratios, the variation over time is often limited. By increasing the covered time period, this could, at least partially, be overcome. However, in our empirical model the size of the time-varying parameter vector $\beta = (\beta_1, \beta_2, ..., \beta_T)'$ grows linearly with the number of time periods. As such, increasing the sample size along the time dimension generally does not lead to an identification improvement. Hence, to mitigate the degree of uncertainty around the estimated path of debt coefficients β_t and to ensure that there is sufficient information present in the data, the number of observations needs to be increased along the cross-sectional dimension. Of course, to benefit from this a homogeneous reaction to debt will be assumed.

However, as shown by Everaert and Jansen (2018), unmodeled slope heterogeneity in the reaction to lagged public debt can lead to the false conclusion that fiscal fatigue is present in the data.[10] Taking into account their results, we will therefore limit our sample to what we believe to be a somewhat homogeneous group of countries.

2.3.2 Dealing with endogeneity issues

When estimating our empirical specification, a potential source of reverse causality or endogeneity is the relationship between pb_{it} and OG_{it}. Fiscal policy, and thus pb_{it}, has an impact on the state of the economy, making the output gap an endogeneous regressor. If not properly accounted for, this might induce (severely) biased estimates. In order to deal with this endogeneity issue, a two-step instrumental variable procedure is employed. The first step constitutes an auxiliary regression that involves regressing OG_{it} on the exogeneous covariates in X_{it} and following Berti et al. (2016) the first and second lag of OG_{it} as instruments for OG_{it}.[11] In the second step and for the remainder of the analysis, in the vector X_{it} we replace OG_{it} by the fitted values of the first-step regression, $\widehat{OG}_{it}$.

2.3.3 Cross-sectional dependencies in the error term

In macroempirical analysis cross-sectional dependencies are more likely to be the rule than the exception, because of strong economic linkages between countries (see Westerlund and Edgerton, 2008). Empirically, this results in significant cross-sectional correlation in the error terms. To deal with this, country and year fixed effects (α_i, δ_t) are employed in equations (2.5) and (2.7). As noted by Eberhardt and Teal (2011), for country and year effects to be efficacious in dealing with cross-sectional correlated error terms, an identical impact of the existing cross-sectional dependence across all countries in the sample needs to be assumed. As our sample is limited to a homogeneous group of core EU countries, we believe this assumption not to be too stringent.

To test whether the included fixed effects adequately control for the presence of cross-sectional dependencies, the average of the country-by-country cross-correlation in the estimated error terms, $\widehat{\epsilon}_{it}$, is calculated. Next to that, we test for the presence of first-order serial correlation in ϵ_{it}, using the Cumby and Huizinga

[10] If there are countries with a weaker reaction to increases in public debt, these countries will eventually end up with a higher debt level. Estimating a homogeneous debt reaction will therefore incorrectly capture this as the presence of fiscal fatigue, while in fact this can be explained by unmodeled slope heterogeneity.

[11] Given the rather high R^2 of the auxiliary regression, the chosen instruments can be regarded as strong. Results from the auxiliary regressions are available upon request.

(1992) test, the reason being that significant autocorrelation would render $d_{i,t-1}$ and its powers endogenous.[12]

2.3.4 Formally testing for time variation in the baseline model

In our empirical specifications, we assume β_t to be time-varying. More precisely, we assume that β_t follows a random walk process. As such, we are estimating a time-varying parameter model. A key issue for the model specification is whether the fiscal policy responsiveness truly varies over time or is constant.

Stated otherwise, the question whether the time variation in β_t is relevant implies testing $\sigma_\eta^2 = 0$ against $\sigma_\eta^2 > 0$, which constitutes a non-regular testing problem as the null hypothesis lies on the boundary of the parameter space. This motivates employing a Bayesian stochastic model specification search (SMSS) algorithm. In a Bayesian setting, each of the potential models is assigned a prior probability and the goal is to derive the posterior probability for each model conditional on the data. The modern approach to Bayesian model selection is to apply Monte Carlo Markov Chain (MCMC) methods by jointly sampling model indicators and parameters. Frühwirth-Schnatter and Wagner (2010) developed this model selection approach for Unobserved Components (UC) models. Their approach relies on a non-centered parameterization of the UC model in which (i) binary stochastic indicators for each of the model components are sampled together with the parameters and (ii) the standard inverse gamma (IG) prior for the variances of innovations to the time-varying components is replaced by a Gaussian prior centered around zero for the standard deviations. In what follows, the exact implementation applied to our baseline model ((2.5)-(2.6)) is outlined.[13] [14]

Non-centered parameterization

As argued by Frühwirth-Schnatter and Wagner (2010), a first piece of information on the hypothesis whether the variance of innovations to a state variable is zero or not can be obtained by considering a non-centered parameterization. This implies rearranging the random walk process for the time-varying parameter β_t, i. e. equation (2.6):

$$\beta_t = \beta_0 + \sigma_\eta \tilde{\beta}_t, \tag{2.10}$$

$$\tilde{\beta}_t = \tilde{\beta}_{t-1} + \tilde{\eta}_t, \quad \tilde{\beta}_0 = 0, \quad \tilde{\eta}_t \sim N(0,1), \tag{2.11}$$

where β_0 is the initial value of β_t if this coefficient varies over time ($\sigma_\eta > 0$), while being the constant value of β_t in case that there is no time variation ($\sigma_\eta = 0$).

[12] As d_{it} is impacted by a contemporaneous shock in the pb_{it}.

[13] This can easily be extended to the more refined specification for β, represented by equations (2.8)-(2.9).

[14] The implementation and description of this approach draws heavily on earlier work from the authors, such as Berger et al. (2016) and Everaert et al. (2017).

A crucial aspect of the non-centered parameterization is that it is not identified: The signs of σ_η and $\tilde{\beta}_t$ can be exchanged without affecting their product. This implies that (i) in the situation where $\sigma_\eta > 0$, the marginal likelihood and therefore the marginal posterior distribution are bimodal with modes $+\sigma_\eta$ and $-\sigma_\eta$ and (ii) when $\sigma_\eta = 0$, the marginal likelihood and posterior will be unimodally centered around zero. As such, allowing for non-identification of σ_η provides useful insights about the degree of time variation governing the debt ratio coefficient.

Parsimonious specification

In the non-centered parameterization, the question whether the fiscal responsiveness to the debt ratio varies over time or not can be expressed as a standard variable selection problem. To this end, a binary indicator that can take the values 0 or 1 is introduced and sampled together with the model's other parameters (see Frühwirth-Schnatter and Wagner, 2010). More precisely, (2.10) becomes

$$\beta_t = \beta_0 + \lambda \sigma_\eta \tilde{\beta}_t, \tag{2.12}$$

where $\lambda \in \{0, 1\}$ is the binary indicator. If $\lambda = 0$, the time-varying component of β_t drops, implying a constant debt ratio coefficient. That is, $\beta_t = \beta_0$ for all t. If $\lambda = 1$, the parameters $\{\tilde{\beta}_1, \tilde{\beta}_2, ..., \tilde{\beta}_T\}$ and σ_η (together representing the time-varying component of β_t) are sampled along with the remaining parameters. We assume a uniform prior distribution for the binary indicator, making both candidate models – the one with a time-varying and the one with a constant debt ratio coefficient – equally likely a priori. Hence, the prior probability is set to $p_0 = 0.5$.

Gaussian priors centered at zero

Our Bayesian estimation approach requires choosing prior distributions for the model parameters. When using the standard inverted Gamma (IG) prior for the variance parameter, the choice of the shape and scale hyperparameters – that define this distribution – has a strong influence on the posterior distribution when the true value of the variance is close to zero (see Frühwirth-Schnatter and Wagner, 2010 and Everaert et al., 2017). As a result, this choice of prior distribution has a tendency to overstate σ_η^2, especially if the true value of σ_η^2 is small.[15] In other words, the actual degree of time variation would be overstated.

When making use of the non-centered parameterization in (2.10)-(2.11) , where σ_η is a regression coefficient, this issue can be resolved. In fact, this allows us to replace the standard IG prior on the variance parameter σ_η^2 by a Gaussian prior centered at zero on σ_η. As the standard deviation σ_η is centered around zero both for $\sigma_\eta^2 > 0$ and $\sigma_\eta^2 = 0$, this makes sense. Frühwirth-Schnatter and Wagner (2010) show that the

[15]More specifically, as the IG distribution does not have probability mass at zero, using it as a prior distribution tends to push the posterior density away from zero.

posterior density of σ_η is much less sensitive to the hyperparameters of the Gaussian distribution and is not pushed away from zero when $\sigma_\eta^2 = 0$.

2.3.5 Estimating the FRF using a Markov Chain Monte Carlo algorithm

For the baseline specification, the system of equations represented by (2.5), (2.12) and (2.11) constitutes a state-space model, where the measurement equation is represented by (2.5) and the state equation for β_t by (2.12) and (2.11). Due to the presence of the time-varying parameter, β_t, and the utilization of the SMSS outlined above, this state-space model is non-standard. We follow Everaert et al. (2017) and employ a Gibbs sampler to estimate the parameters of the state-space model. More specifically, by simulating draws from conditional distributions, thereby breaking the complex estimation problem into easier to handle pieces, a MCMC algorithm is used to obtain approximations of intractable marginal and joint posterior distributions:

For convenience, define $\theta \equiv (\beta_0, \sigma_\eta, \alpha', \delta', \phi, \gamma')'$, $\beta \equiv (\beta_1, \beta_2, ..., \beta_T)'$ and $\tilde{\beta} \equiv (\tilde{\beta}_1, \tilde{\beta}_2, ..., \tilde{\beta}_T)'$, where $\gamma \equiv (\gamma_1, \gamma_2, ...\gamma_k)'$ with k being the number of control variables contained in the predictor matrix X. Next, define a data matrix $Y = (pb, d_{-1}, X)$, where pb, d_{-1} and X contain all observations $i = 1, 2, ..., N$, $t = 1, 2, ..., T$ of $pb_{it}, d_{i,t-1}$ and X_{it}. That is, observations are stacked over cross sections and time periods, with the cross-sectional being the slower index. The resulting MCMC scheme is then given by the following blocks:[16] [17]

1. Sample the binary indicator λ from $p(\lambda|\tilde{\beta}, Y)$, marginalizing over the parameters in θ and σ_ϵ^2, then sample the unrestricted parameters in θ and σ_ϵ^2.

2. Sample the random walk component $\tilde{\beta}$ from $p(\tilde{\beta}|\lambda, \theta, \sigma_\epsilon^2, Y)$.

3. Perform a random sign switch for σ_η and the elements in $\tilde{\beta}$. That is, draw from $\{-1, 1\}$ with equal probability of both outcomes and multiply by σ_η and $\tilde{\beta}$, implying a 50 percent chance of σ_η and $\tilde{\beta}$ being multiplied by (-1). The time-varying parameter vector can then be constructed from its components (based on equation (2.12)).

Given a sufficiently long burn-in phase, the MCMC scheme outlined above produces samples of the parameters that converge to the intractable joint and marginal posterior distributions. We set the total number of Gibbs iterations to 200,000, with a burn-in phase of 80,000. We store every 10th of the remaining 120,000 draws, leaving us with 12,000 retained draws.

[16] Further details of this procedure are laid out in the appendix.

[17] In a separate appendix, more details on the MCMC scheme for the extended model, represented by equation (2.7)-(2.9), are outlined.

2.4 Empirical results

Turning to the empirical results, we first discuss our baseline model in Section 2.4.1. In Section 2.4.2, the model extension is considered, where we allow observed variables to impact on the time-varying path of the primary balance responsiveness.

2.4.1 Baseline specification

The baseline specification refers to the pure random walk model and is represented by equations (2.5), (2.10) and (2.11). As our Bayesian estimation approach requires choosing prior distributions for the model parameters, we will first discuss our prior choices. Then, the results of the SMSS procedure are analyzed, followed by a discussion of the model the SMSS procedure favours, i. e. the parsimonious model.

Prior choices

Summary information on the prior distributions for the unknown parameters is reported in Table 2.1. For the variance σ_ϵ^2 of shocks hitting the primary balance in equation (2.5), an inverse Gamma prior distribution is used, that is $\sigma_\epsilon^2 \sim IG(c_0, C_0)$, where the shape $c_0 = \frac{v_0}{2} NT$ and scale $C_0 = c_0\sigma_0^2$ parameters are calculated from the prior belief σ_0^2 and the prior strength v_0, which are expressed as a fraction of the sample size NT.[18] Our prior belief for σ_ϵ is 1.18, implying that 90% of primary balance shocks lie between -1.99 and 1.99%.[19] Note that the prior is fairly loose as the strength of σ_ϵ^2 is set to $\nu_0 = 0.05$.

For the remaining parameters, Gaussian prior distributions, $N(a_0, V_0)$, are used. Technically, we assume Gaussian parameters and the inverted Gamma distributed regression error variance to jointly follow a dependent Normal-inverted Gamma distribution a priori. This implies the normally distributed parameters to depend on σ_ϵ^2.[20] V_0 then equals $\sigma_2^2 A_0$. When discussing our Gaussian prior choices, first consider the time-varying fiscal responsiveness to the lagged public debt ratio, β_t. For β_0, the prior is given by $\beta_0 \sim N(0, 1.18^2 * 0.32^2)$ (with $\sigma_\epsilon \approx 1.18$), which reflects our belief that if no time variation is present in β_t (i. e. $\sigma_\eta = 0$), then fiscal responsiveness ranges from roughly -0.62 to 0.62. This covers a wide range of parameter values found in the literature.[21] For the standard deviation σ_η of the innovations to the time-varying part in β_t, a Gaussian prior centered at zero is chosen as well ($\sigma_\eta \sim N(0, 1.18^2 * 0.1^2)$). Note that the prior standard deviation of 0.1 implies a very loose prior as it allows that 90% of the innovations to β_t lie between -0.19 and 0.19.

[18] Since the prior is conjugate, $v_0 NT$ can be interpreted as the number of "fictitious" observations used to construct the prior belief σ_0^2 (see also Iseringhausen and Vierke, 2019).

[19] The choice of the prior belief for the standard deviation of $\sigma_0 \approx 1.18$ is based on a standard regression for pb_{it}, estimated with Ordinary Least Squares, where the debt ratio coefficient, β, enters the equation as a constant.

[20] More formally, it holds that $\theta \sim N(a_0, \sigma_\epsilon^2 A_0)$ and $\sigma_\epsilon^2 \sim IG(NT\frac{\nu_0}{2}, NT\frac{\nu_0}{2}$, where $\theta \equiv (\beta_0, \sigma_\eta, \alpha', \delta', \phi, \gamma')'$. Details are provided in appendix 2.A.

[21] Again, we refer the reader to the excellent literature review of Checherita-Westphal and Žd'árek (2017).

TABLE 2.1: Prior choices for the baseline specification

Gaussian priors

$\sim N(a_0, \sigma_\epsilon^2 A_0)$		a_0	$\sqrt{A_0}$	5%	95%
Initial state	β_0	0	0.32	-0.62	0.62
Standard deviation state error	σ_η	0	0.1	-0.19	0.19
AR(1) parameter	ϕ	0.7	0.32	0.06	1.30
Parameters of control variables and fixed effects	$(\gamma', \alpha', \delta')'$	$\underline{0}$	$0.32I$	-0.62	0.62

Inverted Gamma prior

$\sim IG(NT\frac{\nu_0}{2}, NT\frac{\nu_0}{2}\sigma_0^2)$		σ_0	ν_0	5%	95%
Measurement error variance	σ_ϵ	1.18	0.05	0.90	2.36

Notes: For the inverted Gamma prior, we display the prior belief about standard deviation σ_0 instead of the corresponding variance parameter as this is easier to interpret. Likewise, we report $\sqrt{A_0}$ instead of A_0 for the Gaussian priors. For the priors on γ, $\underline{0}$ is a K x 1 vector of zeros, and I is the identity matrix of dimension KxK, with K being the number of control variables and fixed effects.

Simulations of the random walk specification for β_t, based on the prior distribution for σ_η, reveal that the resulting random walk process covers all possible realistic values for the debt ratio coefficient.[22]

For the coefficient on the lagged primary balance, ϕ, a Gaussian prior centered around 0.7 is chosen, roughly in line with estimates found in the literature (see for example Everaert and Jansen, 2018). Although the prior is not centered around 0, we are equally uninformative for ϕ, implying a 90% prior density interval that roughly ranges from 0.06 to 1.30. For the other parameters, i. e. the coefficients on the control variables and the country and time fixed effects, the prior distribution is centered around zero, with the prior standard deviation $\sqrt{A_0}$ being 0.32. As a result, the 90% prior density interval spreads from approximately -0.62 to 0.62, thereby covering a wide range of parameter values found in the literature.

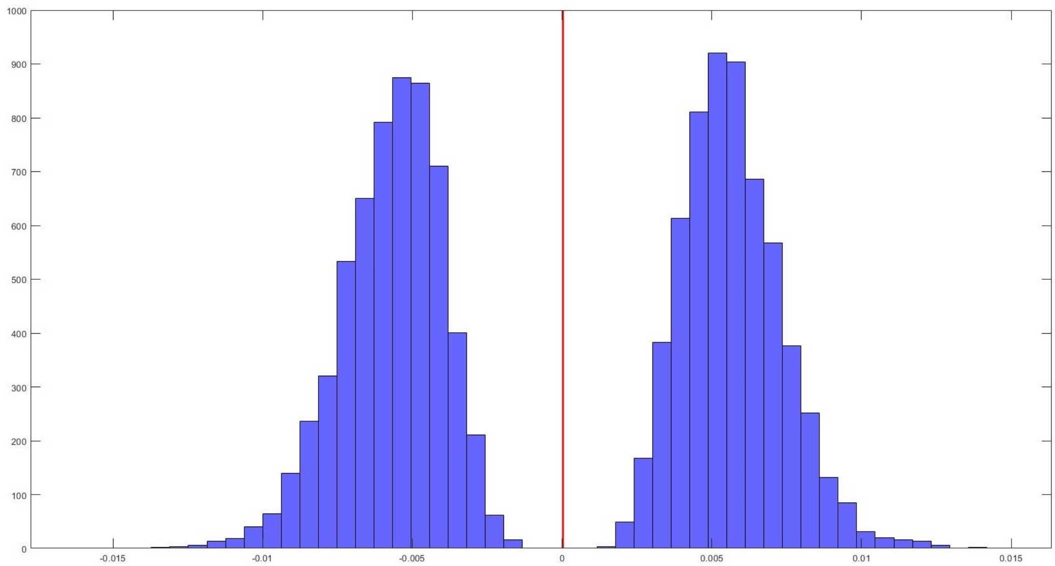
Stochastic model specification search

First, we estimate the unrestricted model, represented by equations (2.5), (2.10) and (2.11). That is, we set the binary indicator λ in (2.12) to 1 to generate a posterior distribution for the standard deviation (σ_η) of the shocks to β_t. If this distribution is bimodal, with low or no probability mass at zero, this can be taken as a first indication of a time-varying primary balance reaction to public debt. Figure 2.1 presents the resulting posterior distribution. Obviously, a clear-cut bimodality in the posterior distribution, with almost no probability mass at zero, is present. This indicates that σ_η^2 indeed appears to be greater than zero.

As a more formal test for the presence of time variation, we next sample the stochastic binary indicator λ along with the unknown parameters. For the binary indicator λ, we choose a Bernoulli prior distribution with a prior probability p_0 of

[22]The simulation results are available upon request.

FIGURE 2.1: Posterior distribution for σ_η, when the binary indicator $\lambda = 1$



being included in the model, i. e. $p(\lambda = 1) = p_0$. As our benchmark, we set p_0 to 0.5, implying indifference between a time-varying and a constant debt ratio coefficient a priori. We further analyze results for the more informative priors $p_0 = 0.25$ and $p_0 = 0.75$. The posterior inclusion probability for the binary indicator is then calculated as the average selection frequency over all iterations of the Gibbs sampler.[23] By looking at the posterior inclusion probability, we obtain valuable information on the question whether time variation is present in the debt ratio coefficient or not.

Table 2.2 displays posterior inclusion probabilities for λ for different prior variances of σ_η and different prior inclusion probabilities. In the baseline case (see Table 2.1), A_0 is set to 0.01 and $p_0 = 0.5$. The posterior inclusion probability of the stochastic binary indicator clearly exceeds 50% and is almost equal to 1. Clearly, based on this result, time variation is present in the fiscal responsiveness to public debt. As a first robustness check, we further consider other values for A_0, given the non-informative prior inclusion probability of λ, $p_0 = 0.5$: For a somewhat stricter prior variance $A_0 = 0.001$, i. e. when placing more weight on values near zero, and a variety of looser prior variances of 0.1, 1 and 10, results are almost equal. In particular, being less informative with respect to σ_η by increasing A_0 hardly brings down the probability of a time-varying fiscal reaction to debt. Thus, even for the very uninformative prior, where $A_0 = 10$, the posterior inclusion probability still amounts to 98.35%. Moreover, the high probability of a time-varying fiscal reaction remains when being more informative with respect to p_0: As expected, when setting $p_0 = 0.75$, posterior inclusion probabilities for λ are even higher. Even when we

[23] In other words, the posterior inclusion probability is the ratio of iterations in which $\lambda = 1$ relative to the absolute number of draws.

TABLE 2.2: Posterior inclusion probabilities for the binary indicator λ

Priors		Posterior
p_0	A_{0,σ_η}	λ
0.25	0.001	0.9992
0.25	0.01	0.9979
0.25	0.1	0.9938
0.25	1	0.9772
0.25	10	0.9399
0.5	0.001	0.9999
0.5	**0.01**	**0.9991**
0.5	0.1	0.9985
0.5	1	0.9925
0.5	10	0.9835
0.75	0.001	0.9999
0.75	0.01	0.9997
0.75	0.1	0.9994
0.75	1	0.9978
0.75	10	0.9928

Note: Results are based upon various prior inclusion probabilities for λ, including a non-informative prior, i. e. $p_0 = 0.5$.

set $p_0 = 0.25$, so that our model favors the finding of no time variation in the fiscal reaction to debt a priori, the posterior inclusion probability clearly exceeds 90% for all four values of A_0. Our evidence regarding time variation in β_t is thus very convincing.[24]

Results parsimonious model

As can be seen from the results in Table 2.2, the SMSS procedure clearly favors a model with the stochastic binary indicator λ set to 1, and where the debt ratio coefficient is allowed to vary over time. In what follows, results of this parsimonious model are discussed.[25]

Table 2.3 shows the estimation results for our set of control variables included in X_{it}. More precisely, as coefficients are sampled along with the other parameters, we report the posterior means as well as the 5th and 95th percentiles of the marginal posterior distributions. The results are broadly in line with the existing literature on FRFs. In particular, we find a positive reaction of the primary balance to an increasing output gap, indicating that fiscal policy is on average countercyclical and thus

[24]While in this paper we focus on a group of what we labeled "core" EU countries, we additionally tested whether a time-varying fiscal response to debt was present in a sample of Southern EU countries, namely Greece, Italy, Portugal and Spain. However, for this sample, the finding of time variation is mixed at best, with a posterior inclusion probability of the stochastic binary indicator being approximately 0.36 in the baseline model. Results are available upon request.

[25]The model is labeled the "parsimonious" model as λ is fixed and thus not sampled along with the other parameters.

can be considered to be an effective stabilization tool.[26] Moreover, inflation is found to have a positive impact on the primary balance, which could be linked to bracket creep effects. Somewhat counterintuitive, according to our results an increase in the interest cost on servicing the public debt would lead to a lower primary balance. As this implies interest costs to rise, one would expect an offsetting impact of the primary balance – as our set of core EU countries is bounded by the nominal deficit rule. Although the mean of the coefficient on the election cycle variable shows that governments tend to increase their spending in election years, the 90% highest posterior density interval clearly encompasses zero. We thus refrain from statements on the effect of a potential political budget cycle. Finally, the posterior distribution for the lagged primary balance coefficient clearly indicates a pronounced sluggishness in the budgeting process.

Note that the bottom of table 2.3 contains information on the possible presence of autocorrelation and cross-sectional dependence in the error term. We follow Everaert and Jansen (2018) in employing the Cumby and Huizinga (1992) test to examine first-order serial correlation in the error terms. An advantage of this test is that it is valid even in the presence of instrumented regressors (and heteroscedasticity). The results suggest that there is no first-order autocorrelation present in the residuals. Related to instrument validity, this implies the lagged endogenous variables are predetermined. Moreover, and equally important, the presence of serial correlation would have rendered d_{t-1} and its powers endogenous. Next to that, the calculated average pairwise correlation in the estimated errors is relatively small, indicating that including country and time fixed effect is sufficient to deal with the possible presence of cross-sectional correlation in the errors.

Of course we are mainly interested in the estimated time-varying path of the debt ratio parameter, β_t, which is displayed in Figure 2.2. The blue line represents the average of the posterior distribution of the time-varying fiscal responsiveness to the public debt ratio, with the shaded area showing the evolution of the 90% highest posterior density interval. The average fiscal responsiveness over time is displayed as a gray line and amounts to 0.013, indicating that Bohn (1998)'s weak sustainability criterion has been fulfilled on average. It should be noted that this value is clearly located near the lower end of the spectrum of debt ratio coefficient estimates found in the literature.

The path shows a weak fiscal reaction to debt in the seventies – with fiscal policy barely responding to changes in debt, as indicated by the 90% highest posterior density interval. However, starting in the beginning of the eighties, the fiscal reaction picked up substantially, with a first peak in the mid-eighties. Subsequently, fiscal responsiveness seemed to stabilize for a certain period, followed by a small increase towards the end of the nineties. The small growth in fiscal responsiveness seems

[26] From a policy point of view, we can make no inference on the relative strength of automatic stabilizers and discretionary fiscal policy. When interested in the relative importance of discretionary fiscal policy as a stabilization tool, one could opt to use the cyclically adjusted primary balance as a dependent variable. This is, however, not the scope of this paper.

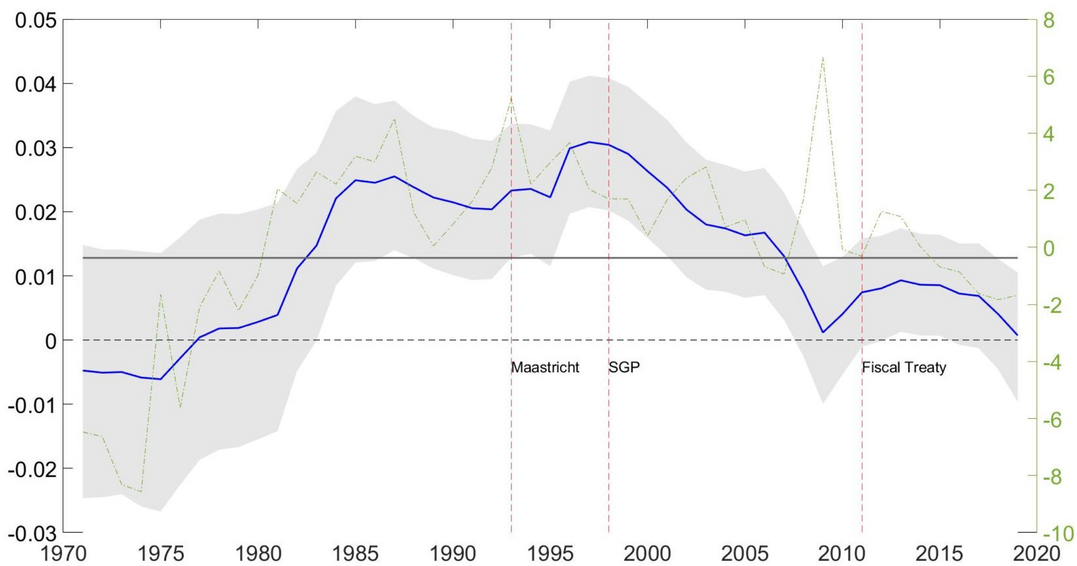
TABLE 2.3: Posterior distribution of main parameters in the parsimonious baseline model

Sample: 1970-2019, 5 core EU countries

Parameter		Posterior mean	5%	95%
Slope parameters				
Output gap	γ_1	0.347	0.224	0.468
Inflation	γ_2	0.142	0.066	0.218
Election cycle	γ_3	-0.134	-0.388	0.122
Implicit interest rate	γ_4	-0.186	-0.275	-0.096
Lagged primary balance	ϕ	0.522	0.435	0.608
Variance parameters				
State error variance	σ_η^2	$3.5e-5$	$1.0e-5$	$7.5e-5$
Measurement error variance	σ_ϵ^2	1.339	1.149	1.550
Residual diagnostics				
Cumby-Huizinga autocorrelation test statistic	0.824	(0.36)		
Average pairwise cross-sectional correlation coefficient	-0.076			

Note: For the Cumby-Huizinga test, the corresponding p-value is put in brackets.

FIGURE 2.2: Time-varying β_t in the baseline specification



Notes: The blue line represents the posterior mean of β_t with the 90% highest posterior density interval as shaded area (left y-axis), while the green line represents the interest-rate growth differential in percent (sample country average, right y-axis). The average debt ratio coefficient is depicted as a gray line. Some milestones of European integration have been added as vertical lines, including the Maastricht Treaty, the Stability and Growth Pact (SGP) as well as the more recently implemented Fiscal Treaty.

to coincide with the period between signing the Maastricht Treaty and the start of the common currency, a possible explanation being that countries needed to fulfill the convergence criteria for adopting the Euro. Afterwards, while the Stability and Growth Pact required continued fiscal efforts from goverments, it seems that fiscal responsiveness dropped signficantly. This preliminary finding confirms the results of Weichenrieder and Zimmer (2014), who relate the drop in fiscal responsiveness after entering the Eurozone to the frequent breaches of the 3% deficit rule, the implicit weakening of the rules and the moral hazard effects from implicit bailout guarantees. The Fiscal Treaty, enforced in 2013, does not appear to have reinforced fiscal prudence, at least according to our anecdotal evidence.

The European Monetary Union has led to a narrowing of spreads among member countries (see amongst others, Turner and Spinelli, 2011), resulting in a downward effect on the IRGD. As such, the downward trend in fiscal responsiveness after Eurozone acceptance could be due to a decreasing IRGD. We therefore also plot the IRGD – averaged over our sample countries – in figure 2.2 in green, with the corresponding values on the right y-axis. Comparing its evolution with the path of the debt ratio coefficient is instructive: In particular, one might argue that the unsubstantial fiscal responsiveness in the seventies might stem from the negative IRGDs at that time. Likewise, the increase in fiscal responsiveness in the eighties as well as the fall after the introduction of the Euro is roughly accompanied by movements of the IRGD in the same direction. A notable exception to this apparent correlation is the spike in IRGDs in the aftermath of the Financial and Economic Crisis of 2007-2008, implied by both lower growth rates and higher refinancing costs in the sample countries as spreads widened. This was due to a sharp increase in public debt ratios and growing concerns from financial markets regarding countries' ability to pay back their debts.

2.4.2 Model extension: Drivers of a time-varying debt coefficient β

As results from the baseline specification show, there is clear evidence of significant time variation in β_t. In a model extension, represented by equations (2.7), (2.8) and (2.9), we therefore model the time-varying fiscal reaction to public debt as a linear combination of a random walk component and a set of covariates (see Section 2.2.2 for more details.), which leads to a country-specific β_{it}. This enables us to shed some light on potential drivers of the observed time-variation in a country's fiscal responsiveness to public debt.

Moreover, by formally testing for time variation in the random walk component by means of the SMSS algorithm, we are able to determine whether there is significant time variation left that cannot be explained by the variables included in G_{it}.

Choice of variables included in G_{it}

As elaborated upon in Section 2.2.2, our analysis focuses on the role of (i) the IRGD and (ii) the level of the public debt ratio in explaining the time variation in fiscal responsiveness to public debt.

- **IRGD:** Focusing on the IRGD is an obvious choice as it plays an essential role in public debt management.[27] If the implicit interest rate on the outstanding amount of debt increases, and thus if debt service costs rise, governments are forced to react stronger to rising public debt ratios in order to prevent an explosive debt path. On the contrary, a higher nominal GDP growth rate tends to lower the debt-to-GDP ratio by increasing the denominator. A lower primary balance will thus be needed to stabilize the public debt ratio. All in all, higher economic growth makes any public debt position more sustainable (ceteris paribus), which justifies a lower fiscal responsiveness to debt. In our empirical analysis, we also allow for an asymmetric government reaction to positive and negative IRGDs. If the IRGD < 0, and nominal interest rates are expected to remain below growth rates for a long time, public debt may have no fiscal costs and only limited welfare costs (see Blanchard, 2019 for an interesting discussion on this topic). On the contrary, a positive and increasing IRGD will lead to a higher primary surplus needed to stabilize or reduce debt. As such, this should lead to an increase in fiscal responsiveness.

- **Lagged debt ratio:** As stated before, an important element in the fiscal austerity debate is whether fiscal responsiveness is impacted by the level of debt itself. By including the lagged debt ratio in G_{it} we actually allow for the presence of non-linearities in the relation between pb_{it} and $d_{i,t-1}$, that are caused by the level of public debt itself. More specifically, taking into account the level of the lagged debt ratio as a covariate in G_{it} implies in fact a parabolic relationship between pb_{it} and $d_{i,t-1}$.

To see this, first consider the non-centered parameterization for equation (2.9),

$$\beta_t^* = \beta_0^* + \lambda \sigma_{\eta^*} \tilde{\beta}_t, \tag{2.13}$$

$$\tilde{\beta}_t = \tilde{\beta}_{t-1} + \tilde{\eta}_t, \quad \tilde{\beta}_0 = 0, \qquad \tilde{\eta}_t \sim N(0,1). \tag{2.14}$$

The extended model is then represented by equations (2.7), (2.8), (2.13) and (2.14), as elaborated upon more thoroughly in appendix 2.B. Now assuming that $d_{i,t-1}$ is the only covariate contained in G_{it}, we can write

$$pb_{it} = \alpha_i + \delta_t + \phi pb_{i,t-1} + \underbrace{(\beta_0^* + \lambda \sigma_{\eta^*} \tilde{\beta}_t + d_{i,t-1}\kappa)}_{\beta_t} d_{i,t-1} + X_{it}\gamma + \epsilon_{it}.$$

Thus, in this case β_0^* captures the constant, linear component of the debt ratio coefficient while κ captures the impact of the quadratic term. A positive but decreasing response of the primary balance to rising debt – and thus a first indication of fiscal fatigue, would show up as $\kappa < 0$.

[27] As argued in Section 2.2.2, we will employ the IRGD at the beginning of period t, $IRGD_{i,t-1}$, as this is the relevant indicator that impacts on the discretionary fiscal policy behavior in t.

Next, and in line with Ghosh et al. (2013), we will allow for a cubic specification to test for the presence of fiscal fatigue in our sample. Analogously to the explanation above, this can be done by including the square of $d_{i,t-1}$ in G_{it}. A negative coefficient on the squared debt term in G_{it} could be seen as evidence for fiscal fatigue.

Finally, recall that an advantage of our approach – compared to Ghosh et al. (2013) and others – is that we also acknowledge that other factors have an impact on β_t. They are captured by the random walk component. As such, possible nonlinearities caused by the actual level of public debt are not the result of ignoring other sources of time variation.

Prior choices

As for the baseline model, our Bayesian estimation approach requires choosing prior distributions for the model parameters in the extended specification. For the parameters already present in the baseline specification, the prior choices are retained. For the new parameter vector κ, a Gaussian prior centered around zero is chosen with the prior standard deviations $\sqrt{A_0}$ being approximately 0.32. Hence, for κ we are also highly uninformative for these parameters as the 90% prior density interval ranges from approximately -0.62 to 0.62.

Stochastic model specification search

As already mentioned, using the SMSS algorithm allows us to make inferences about the importance of including the random walk component in the extended specification for β, i. e. equation (2.8). Results for the SMSS are reported in the upper panel of Table 2.4. Various specifications are estimated, which differ in the variables included in G_{it}. Results show that for all specifications, the posterior inclusion probability $p(\lambda|data)$ of the stochastic binary indicator clearly exceeds 50% and even fluctuates around 90%. This indicates that there is significant time variaton left in the fiscal responsiveness to the public debt ratio that cannot be explained by the variables in G_{it}. However, note that in all specifications, the posterior inclusion probability $p(\lambda|data)$ drops with respect to the baseline model, due to the inclusion of covariates explaining a share of the time variation in β_t. Similarly, the posterior mean of the variance of innovations to the random walk component, σ^2_η, falls for most specifications. Both can be interpreted as preliminary signs that the variables included in G_{it} are indeed explaining part of the observed time variation in fiscal prudence.

Results parsimonious model

The discussion above shows that the SMSS obviously favors a model where the stochastic binary indicator is set to 1 in (2.13). The lower panel of Table 2.4 reports the results for this parsimonious model. More precisely, it presents the posterior mean and the 5th and 95th percentile of the marginal posterior distributions of the impacts (κ) – over different combinations of variables included in G_{it} – on β_{it}.

When looking at the results, one immediately notices that the posterior distributions of the coefficients κ are centered around small numbers. This, however, does not imply that the impact of the included variables is negligible. As the determinants included in G_{it} are trying to explain part of the time variation in β – which is on average equal to 0.013 in our baseline model – it is logical that coefficients are much smaller.[28]

Our empirical results show that the IRGD seems to have no clear impact on a government's fiscal responsiveness. Although the posterior mean has the right positive sign, over the different specifications the 90% highest posterior density interval contains zero (see specifications 1, 5 and 7). However, as the distribution is not centered around zero, this could still be taken as a weak sign that an increase in the IRGD leads to a higher fiscal responsiveness. Moreover, this result could be due to an asymmetric reaction of governments to positive and negative IRGDs. When the IRGD is negative, results clearly show no impact of this determinant on fiscal responsiveness. On the contrary, when the IRGD is positive, an increase in it leads to a rise in fiscal responsiveness. This is expected as a growth in a positive IRGD enlarges the cost for governments of being fiscally irresponsible. This can mainly be seen in specification 2, while in specification 6 and 8 the posterior distribution includes zero but is firmly skewed to the right, with the majority of the probability mass located in the positive area.

Finally, results on potential non-linearities in the fiscal reaction to the public debt ratio show that – at least for our sample – governments tend to respond more when the debt ratio is high. This is clearly confirmed when only including the level of the debt ratio in G_{it} (see specification 3, 5 and 6). When explicitly testing for fiscal fatigue, and thus including a squared debt term in G_{it}, our findings do not provide evidence for the fiscal fatigue proposition of Ghosh et al. (2013) (see specifications 4, 7 and 8). Our results are closer to Everaert and Jansen (2018), who also do not find fiscal fatigue to be a robust characteristic of the fiscal reaction function. However, this does not imply that there is no such debt threshold from where fiscal effort becomes unfeasible or undesirable. But for the debt ratios observed in our homogeneous sample of core EU countries – and when allowing for other determinants driving fiscal responsiveness – we do not find any signs of the fiscal fatigue property.

[28] More precisely, the coefficient vector κ measures the impact on β of a one %-point increase in the corresponding variable.

TABLE 2.4: Results extended model

Sample: 1970-2019, 5 core EU countries

	Specification 1			Specification 2			Specification 3			Specification 4		
Stochastic model specification search												
	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
$\sigma^2_{\tilde{\eta}}$	$3.4e{-}5$	$9.0e{-}6$	$7.6e{-}5$	$3.9e{-}5$	$1.1e{-}5$	$8.5e{-}5$	$2.9e{-}5$	$6.5e{-}6$	$6.7e{-}5$	$3.0e{-}5$	$6.8e{-}6$	$6.8e{-}5$
$p(\lambda\|data)$		0.978			0.988			0.891			0.887	
Results posterior distributions for κ in the parsimonious model												
	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
IRGD	0.0007	-0.0005	0.002	-	-	-	-	-	-	-	-	-
IRGD<0	-	-	-	-0.0008	-0.0036	0.0022	-	-	-	-	-	-
IRGD>0	-	-	-	0.001	0.000	0.003	-	-	-	-	-	-
Debt	-	-	-	-	-	-	0.0002	$4.2e{-}5$	0.0003	-0.0005	-0.0013	0.0003
Debt^2	-	-	-	-	-	-	-	-	-	$3.8e{-}6$	$-3.7e{-}7$	$7.9e{-}6$

Notes: $p(\lambda|data)$ represents the posterior inclusion probability for the random walk component of the non-centered parameterization. Further, posterior moments are displayed for the variance of the random walk component and for the set of parameters κ in the parsimonious model.

Table 2.4, continued

	Specification 5			Specification 6			Specification 7			Specification 8		
Stochastic model specification search												
	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
$\sigma^2_{\tilde{\eta}}$	$2.9e{-}5$	$6.1e{-}6$	$6.6e{-}5$	$3.5e{-}5$	$8.4e{-}6$	$7.9e{-}5$	$2.9e{-}5$	$6.4e{-}6$	$6.8e{-}5$	$3.6e{-}5$	$9.0e{-}6$	$8.0e{-}5$
$p(\lambda \mid data)$		0.820			0.911			0.805			0.895	
Results posterior distributions for κ in the parsimonious model												
	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
IRGD	0.0002	-0.0011	0.0015	-	-	-	0.0002	-0.0011	0.0015	-	-	-
IRGD<0	-	-	-	-0.0019	-0.005	0.001	-	-	-	-0.0018	-0.0048	0.0012
IRGD>0	-	-	-	0.0008	-0.0008	0.0023	-	-	-	0.0007	- 0.0008	0.0022
Debt	0.0002	$2.5e{-}5$	0.0003	0.0002	4.6e-5	0.0004	-0.0005	-0.0014	0.0003	-0.0005	-0.0013	0.0003
Debt^2	-	-	-	-	-	-	$3.8e{-}6$	-3.1$e{-}7$	$7.9e{-}6$	$3.6e{-}6$	$-5.4e{-}7$	$7.7e{-}6$

Notes: $p(\lambda|data)$ represents the posterior inclusion probability for the random walk component of the non-centered parameterization. Further, posterior moments are displayed for the variance of the random walk component and for the set of parameters κ in the parsimonious model.

2.5 Conclusion

The fiscal policy response to the COVID-19 crisis has put severe pressure on public finances in the EU. Against the backdrop of low expected potential economic growth and sharply rising age-related public expenditures, this has revived the debate on the sustainability of public finances. Estimating FRFs and empirically analyzing whether countries react to a growing public debt ratio by tightening the fiscal policy stance can shed light on a country's degree of fiscal prudence.

The FRF literatue has well recognized the importance of non-linearities for correctly specifying the FRF. A key issue is whether the degree of fiscal responsiveness changes with the level of debt. Specifically, as introduced by Ghosh et al. (2013), the hypothesis of fiscal fatigue has been tested, implying that at a certain debt level the fiscal response starts to weaken and even decreases.

In this paper, we formally test for the presence and potential sources of non-linearities by allowing for a time-varying fiscal responsiveness to debt. This approach is related to commonly used FRF specifications that embed the fiscal fatigue proposition, but is more flexible as it allows non-linearities in the fiscal reaction to debt to arise stochastically by means of a time-varying parameter model.

Having employed a Bayesian SMSS testing procedure to formally test for the presence of time variaton in the responsiveness in the primary balance to the gross public debt ratio, we find strong evidence for time-variation in the FRF over the last 50 years. Governments' fiscal stance to debt exhibits smooth but significant variation over time and thus confirms the necessity of a non-linear model.

In a model extension, we explicitly try to make inferences about potential driving forces of the time varyings fiscal responsiveness. As such, we are able to test for the presence of fiscal fatigue in a stochastic way, i. e. acknowledging the potential presence of other sources of time variation.

Our results provide preliminary evidence that the fiscal response to debt seems to be partly explained by changes in the IRGD, at least when the IRGD is positive. In that case, an increase in the IRGD reinforces the cost of being fiscally irresponsible. Governments will therefore react by tightening their fiscal policy stance.

When allowing for non-linearities caused by the level of public debt, our model does not provide robust evidence of the fiscal fatigue proposition of Ghosh et al. (2013). On the contrary, the results indicate that – for our sample – governments tend to increase fiscal responsiveness when the debt ratio increases. As such, these results are more in line with the findings of Everaert and Jansen (2018), who also do not find fiscal fatigue to be a robust characteristic of the FRF. However, this does not imply that no such debt threshold exists from which the fiscal effort becomes unfeasible or undesirable. It is just not observed in our sample of public debt ratios.

Our findings further indicate that a significant fraction of the time variation governing the fiscal reaction coefficient is not explained by our set of predictors. Future research on potential other sources of the observed time variation could therefore be very clarifying and help in explaining countries' fiscal stances to debt. More precisely, it could be interesting to look explicitly into the relevance of financial market pressure and the role of political economy determinants, such as the political orientation of governments.

Given their prominence in stochastic DSA, a correctly specified FRF is of utmost importance. In our analysis, we propose a careful assessment of whether potential parameter instability should be accounted for in the sample of interest. Our results clearly indicate that time-varying FRFs appear to be an adequate choice.

Bibliography

Berger, T., Everaert, G., and Vierke, H. (2016). Testing for time variation in an unobserved components model for the u.s. economy. *Journal of Economic Dynamics and Control*, 69:179 – 208.

Berti, K., Colesnic, E., Desponts, C., Pamies, S., and Sail, E. (2016). Fiscal reaction functions for European Union countries. European Economy - Discussion Papers 2015 - 028, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Blanchard, O. (2019). Public debt and low interest rates. *American Economic Review*, 109(4):1197–1229.

Bohn, H. (1998). The behavior of U. S. public debt and deficits. *The Quarterly Journal of Economics*, 113(3):949–963.

Burger, P., Stuart, I., Jooste, C., and Cuevas, A. (2011). Fiscal sustainability and the fiscal reaction function for south africa. *IMF Working Papers*, pages 1–27.

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

Checherita-Westphal, C. and Žd'árek, V. (2017). Fiscal reaction function and fiscal fatigue: Evidence for the euro area. ECB Working Paper 2036, European Central Bank.

Cumby, R. E. and Huizinga, J. (1992). Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions. *Econometrica*, 60(1):185–195.

Debrun, X., Moulin, L., Turrini, A., i Casals, J. A., Kumar, M. S., Drazen, A., and Fuest, C. (2008). Tied to the mast? national fiscal rules in the european union [with discussion]. *Economic Policy*, 23(54):297–362.

Debrun, X., Ostry, J. D., Willems, T., and Wyplosz, C. (2019). Public debt sustainability. CEPR Discussion Papers 14010, C.E.P.R. Discussion Papers.

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

Eberhardt, M. and Teal, F. (2011). Econometrics for grumblers: A new look at the literature on cross-country growth empirics. *Journal of Economic Surveys*, 25(1):109–155.

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.

Fournier, J.-M. and Fall, F. (2017). Limits to government debt sustainability in oecd countries. *Economic Modelling*, 66:30–41.

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.

Havik, K., Morrow, K. M., Orlandi, F., Planas, C., Raciborski, R., Roeger, W., Rossi, A., Thum-Thysen, A., and Vandermeulen, V. (2014). The production function methodology for calculating potential growth rates & output gaps. European Economy - Economic Papers 2008 - 2015 535, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Iseringhausen, M. and Vierke, H. (2019). What drives output volatility? the role of demographics and government size revisited. *Oxford Bulletin of Economics and Statistics*, 81(4):849–867.

Koop, G. M. (2003). *Bayesian econometrics*. John Wiley & Sons Inc.

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.

Saez, E. (2003). The effect of marginal tax rates on income: A panel study of 'bracket creep'. *Journal of Public Economics*, 87(5):1231 – 1258.

Turner, D. and Spinelli, F. (2011). Explaining the interest-rate-growth differential underlying government debt dynamics. *OECD Economics Department Working Papers, No. 919*.

Weichenrieder, A. J. and Zimmer, J. (2014). Euro membership and fiscal reaction functions. *International Tax and Public Finance*, 21(4):598–613.

Westerlund, J. and Edgerton, D. L. (2008). A simple test for cointegration in dependent panels with structural breaks. *Oxford Bulletin of Economics and Statistics*, 70(5):665–704.

Appendix

2.A Gibbs sampling procedure pure random walk model

In this section, we provide details on the Gibbs sampling algorithm for the "pure random walk model". The full model in this case consists of the equations (2.5), (2.12) and (2.11), restated here for convenience (with the slight notational difference that the regressor matrices corresponding to the fixed effects are now contained in X, the parameters $\alpha \equiv (\alpha_1, \alpha_2, ..., \alpha_N)'$ and $\delta \equiv (\delta_2, \delta_3 ..., \delta_T)'$ thus contained in γ):

$$
\begin{aligned}
pb_{it} =& \phi pb_{i,t-1} + \beta_t d_{i,t-1} + X_{it}\gamma + \epsilon_{it}, && && \epsilon_{it} \sim N(0, \sigma_\epsilon^2), && (2.15)\\
\beta_t =& \beta_0 + \lambda \sigma_\eta \tilde{\beta}_t, && && && (2.16)\\
\tilde{\beta}_t =& \tilde{\beta}_{t-1} + \tilde{\eta}_t, && \tilde{\beta}_0 = 0, && \tilde{\eta}_t \sim N(0,1). && (2.17)
\end{aligned}
$$

In what follows we provide details on the MCMC algorithm employed to jointly sample the time-varying parameter vector β, the hyperparameters collected in θ and σ_ϵ^2 and the stochastic binary indicator λ. The outlined procedure is based on Frühwirth-Schnatter and Wagner (2010) and Berger et al. (2016).

2.A.1 Sampling the stochastic binary indicator and the hyperparameters

In this block, we sample the stochastic binary indicator λ and the hyperparameters, collected in θ and σ_ϵ^2. For notational convenience, define a general regression model

$$
y = \chi^m \theta^m + e, \quad e \sim N(0, \Sigma), \tag{2.18}
$$

where y is the dependent variable vector and χ is an unrestricted predictor matrix corresponding to the parameter vector $\theta \equiv (\beta_0, \sigma_\eta, \phi, \gamma')'$. 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. Correspondingly, χ^m and θ^m are the restricted predictor matrix and parameter vector, where σ_η and its associated predictor vector are excluded from θ and χ if the binary indicator λ is zero. The covariance matrix of the error term e is a diagonal matrix simply given by $\Sigma = diag\left(\sigma_\epsilon^2 I_{NT}\right)$, where I_{NT} is the identity matrix of dimension NT with N and T being the numbers of cross-sectional and time units in the sample, respectively. σ_ϵ^2 is a scalar. Thus, we assume homoscedasticity.

Note that simply drawing from $p(\lambda|\theta, \sigma_\epsilon^2, \tilde{\beta}, y, \chi)$ and $p(\theta, \sigma_\epsilon^2|\lambda, \tilde{\beta}, y, \chi)$ does not yield an irreducible Markov chain as a draw of $\lambda = 0$ implies that σ_η will also be zero, which leads to the Markov chain having absorbing states. We follow Frühwirth-Schnatter and Wagner (2010) in resolving this by marginalizing over the parameters in θ and σ_ϵ^2 when drawing λ and subsequently sampling from $p(\theta, \sigma_\epsilon^2|\lambda, \tilde{\beta}, y, \chi)$.

The posterior distribution of λ is obtained from Bayes' rule:

$$p(\lambda|\tilde{\beta}, y, \chi) \propto f(y|\lambda, \tilde{\beta}, \chi)p(\lambda), \tag{2.19}$$

where $f(y|\lambda, \tilde{\beta}, \chi)$ is the marginal likelihood of the regression model in (2.18), having integrated out θ and σ_ϵ^2, and $p(\lambda)$ is the prior distribution of λ.

Given homoscedasticity, a dependent Normal-inverted Gamma prior with $\theta^m \sim N(a_0^m, A_0^m\sigma_\epsilon^2)$ and $\sigma_\epsilon^2 \sim IG(c_0, C_0)$, with c_0 and C_0 being the shape and scale parameters of the prior distribution for the measurement error variance, is conjugate, implying the closed form solution of the marginal likelihood[29]

$$f(y|\lambda, \tilde{\beta}, \chi) \propto \frac{|A_T^m|^{0.5}}{|A_0^m|^{0.5}} \frac{\Gamma(c_T)C_0^{c_0}}{\Gamma(c_0)\,(C_T^m)^{c_T}}, \tag{2.20}$$

where

$$a_T^m = A_T^m\left((\chi^m)'y + (A_0^m)^{-1}a_0^m\right), \tag{2.21}$$

$$A_T^m = \left((\chi^m)'\chi^m + (A_0^m)^{-1}\right)^{-1}, \tag{2.22}$$

$$c_T = c_0 + \frac{NT}{2}, \tag{2.23}$$

$$C_T = C_0 + 0.5\left(y'y + (a_0^m)'\,(A_0^m)^{-1}\,a_0^m - (a_T^m)'\,(A_T^m)^{-1}\,a_T^m\right). \tag{2.24}$$

The above can then be applied to the state-space model in equations (2.15), (2.16) and (2.17).

Inserting (2.16) into (2.15) yields

$$pb_{it} = \phi pb_{i,t-1} + \beta_0 d_{i,t-1} + \lambda\sigma_\eta\tilde{\beta}_t d_{i,t-1} + X_{it}\gamma + \epsilon_{it}, \qquad \epsilon_{it} \sim N(0, \sigma_\epsilon^2), \tag{2.25}$$

which can be written as

$$\underbrace{pb_{it}}_{y_{it}} = \underbrace{\begin{bmatrix} d_{i,t-1} & \lambda\tilde{\beta}_t d_{i,t-1} & pb_{i,t-1} & X_{it} \end{bmatrix}}_{\chi_{it}^m} \underbrace{\begin{bmatrix} \beta_0 \\ \sigma_\eta \\ \phi \\ \gamma \end{bmatrix}}_{\theta^m} + \epsilon_{it}. \tag{2.26}$$

[29] Note that we follow Berger et al. (2016) in employing a *dependent Normal-inverted Gamma prior* due to the assumption of homoscedasticity. Hence, the prior variance parameters $V_0 \equiv \sigma_\epsilon^2 A_0$ cannot simply be interpreted as the prior covariance matrix of the normally distributed parameters, as V_0 depends on σ_ϵ^2. Details can be found in Koop (2003), who uses precision instead of variance parameters, however (and therefore works with Normal-Gamma, not Normal-inverted Gamma distributions).

Note that the second elements of χ_{it}^m and θ^m are excluded (set to zero) if $\lambda = 0$, while for $\lambda = 1$, σ_η is sampled along with the other parameters in θ. The marginal likelihood $f(y|\lambda, \beta)$ is then given by (2.20), and the stochastic binary indicator λ can be sampled from the Bernoulli distribution:

$$p(\lambda = 1|\beta, y, \chi) = \frac{f(\lambda = 1|\beta, y)}{f(\lambda = 0|\beta, y) + f(\lambda = 1|\beta, y)} \tag{2.27}$$

Given λ, θ^m and σ_ϵ^2 can then be sampled jointly from $\theta^m, \sigma_\epsilon^2 \sim NIG(a_T^m, A_T^m, c_T, C_T)$, where the posterior moments are given by (2.21), (2.22), (2.23) and (2.24).

2.A.2 Sampling the time-varying parameter

In this block, we employ the forward-filtering backward-sampling procedure of Carter and Kohn (1994) to sample the time-varying component $\tilde{\beta}$ given θ, σ_ϵ^2 and λ. Our conditional linear Gaussian state-space model is given by:

$$y_t = H_t^m s_t^m + e_t, \qquad\qquad e_t \sim MN(0_N, R), \tag{2.28}$$

$$s_t = F s_{t-1} + K_t v_t, \qquad s_0 \sim N(b_0, V_0), \qquad v_t \sim N(0, Q), \tag{2.29}$$

where y_t is an N x 1 vector of observations and H_t^m is the restricted version of the predictor matrix, with s_t^m being the corresponding time-varying parameter vector, for which $H_t^m = H_t$ and $s_t^m = s_t$ in the unrestricted case. 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, we refer to Durbin and Koopman (2012).

We can then employ the Kalman filter 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).

Rearranging terms in equation (2.25) and restating the state equation (2.17) yields the unrestricted conditional state-space model for $\tilde{\beta}_t$:

$$\overbrace{pb_{it} - \phi pb_{i,t-1} - d_{i,t-1}\beta_0 - X_{it}\gamma}^{y_{it}} = \overbrace{d_{i,t-1}\lambda\sigma_\eta}^{H_t^m} \overbrace{\tilde{\beta}_t}^{s_t^m} + \overbrace{\epsilon_{it}}^{e_{it}}, \qquad \epsilon_{it} \sim N(0, \overbrace{\sigma_\epsilon^2}^{R}), \tag{2.30}$$

$$\underbrace{\tilde{\beta}_t}_{s_t} = \underbrace{1}_{F} \underbrace{\tilde{\beta}_{t-1}}_{s_{t-1}} + \underbrace{1}_{K_t} \underbrace{\tilde{\eta}_t}_{v_t}, \qquad \tilde{\eta}_t \sim N(0, \underbrace{1}_{Q}). \tag{2.31}$$

Notice that s_t^m is a scalar as we assume the time-varying parameter to be homogeneous across countries, as outlined above. Stacking observations over $i = 1, 2, ..., N$, this can be written as

$$\overbrace{\begin{bmatrix} pb_{1t} - pb_{1,t-1}\phi - d_{1,t-1}\beta_0 - X_{1t}\gamma \\ \vdots \\ pb_{Nt} - pb_{N,t-1}\phi - d_{N,t-1}\beta_0 - X_{Nt}\gamma \end{bmatrix}}^{y_t} = \overbrace{\begin{bmatrix} d_{1,t-1}\lambda\sigma_\eta \\ \vdots \\ d_{N,t-1}\lambda\sigma_\eta \end{bmatrix}}^{H_t^m} \overbrace{\tilde{\beta}_t}^{s_t^m} + \overbrace{\begin{bmatrix} \epsilon_{1t} \\ \vdots \\ \epsilon_{Nt} \end{bmatrix}}^{e_t}, \tag{2.32}$$

$$\overbrace{\begin{bmatrix} \epsilon_{1t} \\ \vdots \\ \epsilon_{Nt} \end{bmatrix}}^{e_t} \sim \left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \overbrace{\sigma_\epsilon^2 \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}}^{R} \right), \tag{2.33}$$

$$\underbrace{\tilde{\beta}_t}_{s_t} = \underbrace{1}_{F} \underbrace{\tilde{\beta}_{t-1}}_{s_{t-1}} + \underbrace{1}_{K_t} \underbrace{\tilde{\eta}_t}_{v_t}, \tag{2.34}$$

$$\underbrace{\tilde{\eta}_t}_{v_t} \sim N(0, \underbrace{1}_{Q}), \tag{2.35}$$

The time-varying component $\tilde{\beta}_t$ is initialized with mean and variance $b_0 = 0$ and $P_0 = 0.00001$. By doing so, we ensure that the time-varying parameter β_t is initialized with its first value, 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, we follow the univariate treatment of the multivariate series approach 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.

In the restricted model, that is, for $\lambda = 0$, χ^m and s^m are empty. Thus, no forward-filtering and backward-sampling is applied. In this case, $\tilde{\beta}_t$ is sampled directly from its prior, that is, from (2.17).

Lastly, given its components β_0, σ_η and $\tilde{\beta}$, the time-varying parameter vector β can be constructed from (2.16).

2.B Gibbs sampling procedure extended model

In this section, we lay out the Gibbs sampling procedure for the "extended model", where the time-varying parameter equation contains a set of covariates G. Formally, this entails employing the non-centered parameterization for $\tilde{\beta}_t$ and introducing a stochastic binary indicator, sampled along with the other parameters, analogous to the model in equations (2.5), (2.12) and (2.11) referred to in section 2.3.5 and appendix 2.A (we will refer to this model as the "simpler model" below). This implies that the extended model is comprised of the following set of equations (analogous to the notation in appendix 2.A):[30]

$$pb_{it} = pb_{i,t-1}\phi + d_{i,t-1}\beta_{it} + X_{it}\gamma + \epsilon_{it}, \qquad \epsilon_{it} \sim N(0,\sigma_\epsilon^2), \tag{2.36}$$

$$\beta_{it} = \beta_t^* + G_{it}\kappa, \tag{2.37}$$

$$\beta_t^* = \beta_0^* + \lambda\sigma_{\eta^*}\tilde{\beta}_t, \tag{2.38}$$

$$\tilde{\beta}_t = \tilde{\beta}_{t-1} + \tilde{\eta}_t, \qquad \tilde{\eta}_t \sim N(0,1), \tag{2.39}$$

where (2.9) has been replaced by (2.38) and (2.39) and G is a set of covariates potentially driving the time variation in the fiscal reaction to debt. By sampling the stochastic binary indicator λ along with the other parameters, we obtain useful information as to whether the time-varying component $\sigma_{\eta^*}\tilde{\beta}_t$ contains any further information beyond that in the covariates G.

To simplify notation, we now include the parameters of the covariates G, that is κ, in the parameter vector θ, so that $\theta \equiv (\beta_0, \sigma_\eta, \phi, \gamma', \kappa')'$, with $\kappa \equiv (\kappa_1, \kappa_2, ..., \kappa_s)$, where s is the number of explanatory variables included in the state equation, and $\tilde{\beta} \equiv (\tilde{\beta}_1, \tilde{\beta}_2, ..., \tilde{\beta}_T)'$. Analogous to the simpler model, the MCMC scheme splits the estimation problem into three blocks where the parameters are drawn from conditional distributions:

1. Sample the binary indicator λ from $p(\lambda|\tilde{\beta}, Y)$, marginalizing over the parameters in θ and σ_ϵ^2, then sample the unrestricted parameters in θ and σ_ϵ^2.

2. Sample the time-varying parameter vector $\tilde{\beta}$ from $p(\tilde{\beta}|\lambda, \theta, \sigma_\epsilon^2, Y)$.

3. Perform a random sign switch for σ_{η^*} and the elements in $\tilde{\beta}$. That is, draw from $\{-1, 1\}$ with equal probability of both outcomes and multiply by σ_{η^*} and $\tilde{\beta}$, implying a 50 percent chance of σ_{η^*} and $\tilde{\beta}$ being multiplied by (-1). β^* and β can then be constructed from their components.

In what follows, we lay out this MCMC scheme in more detail. Given the similarity of this approach with that of the simpler model, the following sections are mainly concerned with elaborating on the differences between the two.

[30] In addition to the sources mentioned in appendix 2.A, this section draws from Iseringhausen and Vierke (2019).

2.B.1 Sampling the stochastic binary indicator and the hyperparameters

Analogously to the procedure in the pure random walk case, insert (2.37) in (2.36), using the expression for β_t^* in (2.38) to obtain

$$\underbrace{pb_{it}}_{y_{it}} = \underbrace{\begin{bmatrix} d_{i,t-1} & d_{i,t-1}\lambda\tilde{\beta}_t & d_{i,t-1}G_{it} & pb_{i,t-1} & X_{it} \end{bmatrix}}_{X_{it}^m} \underbrace{\begin{bmatrix} \beta_0^* \\ \sigma_{\eta^*} \\ \kappa \\ \phi \\ \gamma \end{bmatrix}}_{\theta^m} + \epsilon_{it}. \tag{2.40}$$

Thus, θ^m now additionally contains the parameters of the covariates in the state equation κ, and the sampling scheme laid out in section 2.A.1 can be employed.

2.B.2 Sampling the time-varying parameter

As before, in this block we set up the conditional state-space model for β_t^*:

$$\overbrace{pb_{it} - \phi pb_{i,t-1} - d_{i,t-1}\beta_0^* - d_{i,t-1}G_{it}\kappa - X_{it}\gamma}^{y_{it}} = \overbrace{d_{i,t-1}\lambda\sigma_{\eta^*}}^{H_t^m} \overbrace{\tilde{\beta}_t}^{s_t^m} + \overbrace{\epsilon_{it}}^{e_{it}}, \quad \epsilon_{it} \sim N(0, \overbrace{\sigma_\epsilon^2}^{R}), \tag{2.41}$$

$$\underbrace{\tilde{\beta}_t}_{s_t} = \underbrace{1}_{F} \underbrace{\tilde{\beta}_{t-1}}_{s_{t-1}} + \underbrace{1}_{K_t} \underbrace{\tilde{\eta}_t}_{v_t}, \quad \tilde{\eta}_t \sim N(0, \underbrace{1}_{Q}), \tag{2.42}$$

for each $i = 1, 2, ..., N$ and $t = 1, 2, ...T$. Given this conditional state-space model, the time-varying parameter $\tilde{\beta}_t$ is sampled just as in the baseline model in section 2.A.2. Lastly, given β_0^*, σ_{η^*} and $\tilde{\beta}$, β^* and β can be constructed from their components.

As for the simpler model, we set the total number of Gibbs iterations to 200,000, with a burn-in phase of 80,000, keeping every 10th draw of the remaining 120,000, which leaves us with 12,000 retained draws.

Chapter 3

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

accepted for publication in *Applied Economics*

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

3.1 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 3.2 elaborates on the basics of SDSA and lays out the benchmark SDSA model. In section 3.3, data, priors and the results are presented. Section 3.4 concludes.

3.2 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).

3.2.1 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:

$$debt_t = \frac{1+i_t}{1+g_t} debt_{t-1} - pb_t, \qquad t = 1, 2, ..., T, \tag{3.1}$$

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 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.

[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.

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.

3.2.2 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

$$pb_t = \alpha + debt_{t-1}\beta_1 + X_t\gamma + \epsilon_t, \tag{3.2}$$

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.

[2]Some additional information on the "fixed coefficient approach" can be found in appendix 3.B.4.
[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).

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 (3.2), I follow Berger et al. (2021) in estimating a FRF of the form

$$pb_{it} = \alpha_i + H_{it}\beta_t + X_{it}\gamma + \epsilon_{it}, \tag{3.3}$$

$$\epsilon_{it} = \mu_t + \rho\epsilon_{i,t-1} + u_{it}, \quad u_{it} \sim N(0, \sigma^2_{u_i}), \quad i = 1, 2, ..., N, \quad t = 1, 2, ..., T, \tag{3.4}$$

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 (3.2):

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 Žd'árek, 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 (3.3) 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 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.

[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.

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]

$$\beta_t = \beta_0 + \sigma_\eta \tilde{\beta}_t, \tag{3.5}$$

$$\tilde{\beta}_t = \tilde{\beta}_{t-1} + \tilde{\eta}_t, \qquad \tilde{\beta}_0 = 0, \qquad \tilde{\eta}_t \sim N(0,1). \tag{3.6}$$

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 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 (3.3), (3.4), (3.5) and (3.6) 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

[6]Note that the non-centered parameterization (NCP) is simply a reparameterization of the random walk process.
[7]A detailed version of the algorithm can be found in the appendix.

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.

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.

3.2.3 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

$$y_t = \phi_{1,t}y_{t-1} + \phi_{2,t}y_{t-2} + ... + \phi_{p,t}y_{t-p} + u_t, \qquad u_t \sim N(0, \Sigma), \tag{3.7}$$

$$\Phi_t = \Phi_{t-1} + e_t, \qquad e_t \sim N(0, Q), \tag{3.8}$$

$t = \{1, 2, ..., T_q\}$, 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 (3.8).

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 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.

3.2.4 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

[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 3.B.2.

(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 (3.7) and (3.8)) 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 (3.3), (3.4), (3.5) and (3.6) 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 (3.1) 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.

3.3 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

[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.

time variation in the coefficients β_t is governed 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).

3.3.1 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 3.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 3.A.

3.3.2 Priors

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

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.

[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.

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 3.1. 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 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 Žd'árek, 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 3.1. 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.

[13]Note that these fixed coefficients are simply the estimates of the fixed coefficient model ("fixed model"), presented in appendix 3.B.4.

TABLE 3.1: 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

$\sim IG(T\frac{\nu_{0,i}}{2}, T\frac{\nu_{0,i}}{2}\sigma^2_{0,i})$		$\sigma_{0,i}$	$\nu_{0,i}$	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.

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 3.1 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).

BVAR priors

This section outlines the priors for the BVAR in equations (3.7) and (3.8).[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 (3.7), 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 (3.7) 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.

[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.

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, 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.

3.3.3 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.

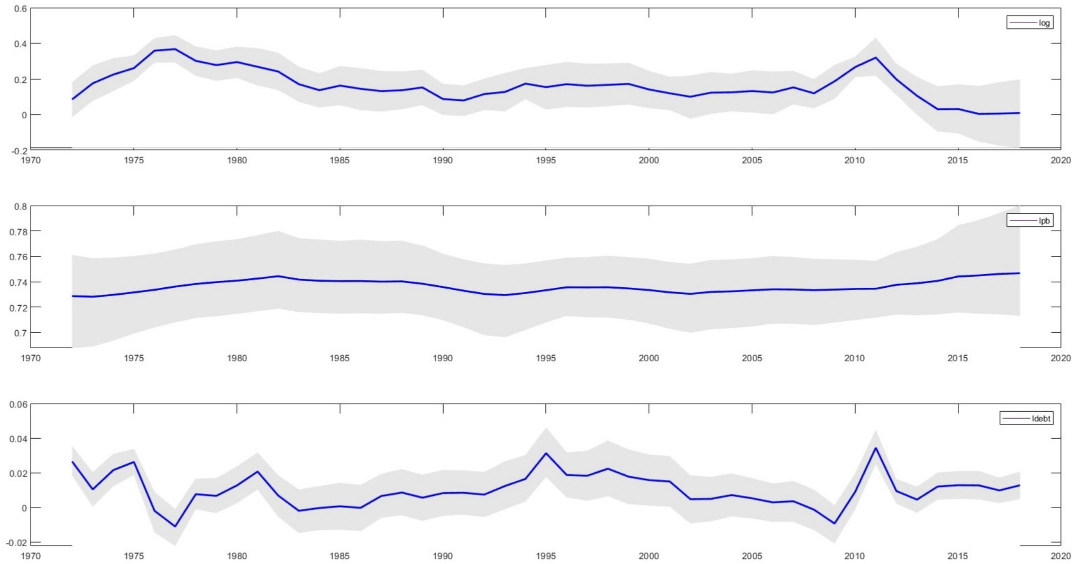
Figure 3.1 displays the paths of the three time-varying parameters, including their 90% credible sets. The top-most panel displays the evolution of the coefficient on the output gap. First, note that there appears to be a certain degree of time variation present. Such non-linearities in the fiscal response 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 Žd'árek, 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 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

FIGURE 3.1: 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.

the set of covariates contained in X and H in (3.3) 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 3.2. Note that in the baseline specification, X and γ in (3.3) are empty, since the coefficients of the lagged primary balance, the lagged debt ratio and the (instrumented) output gap are all modeled in a time-varying manner. Thus, table 3.2 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.

3.3.4 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 3.2.1 and appendix 3.B.4

TABLE 3.2: 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.

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 3.3 - 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 3.3 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 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 3.4 and 3.5. 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 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.

[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 3.3: 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 3.4: 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.

TABLE 3.5: 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.

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

$$z_{it} = \left[e(h)^A_{it}\right]^2 - \left[e(h)^B_{it}\right]^2, \tag{3.9}$$

$i = 1, 2, ..., N$, $t = 1, 2, ..., T$, 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

[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.

[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 3.6: 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.

the null hypothesis that $\alpha_i = 0$ for all $i = 1, 2, ..., N$ in

$$z_{it} = \alpha_i + \epsilon_{it}, \quad \epsilon_{it} \sim IID(0, \sigma_i^2), \tag{3.10}$$

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

$$\overline{DM} = \frac{\bar{z}}{\sqrt{V(\bar{z})}} \sim N(0,1), \tag{3.11}$$

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 3.6 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.

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.

3.3.5 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 3.1). 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 (3.3)). 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 3.7 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 3.1, 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 3.8, 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 3.7: 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 $pb_{it} - pb^F_{it} = \alpha + u_{itH}$ is tested, where pb_{itH} 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 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 3.8: 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.

3.4 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.

Bibliography

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, European Central Bank.

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 Žd'árek, V. (2017). Fiscal reaction function and fiscal fatigue: Evidence for the euro area. ECB Working Paper 2036, European Central Bank.

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?

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.

Appendix

3.A Data

This section provides details about the data used in this paper. Table 3.A.1 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 3.A.2. 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 3.A.3.

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.

TABLE 3.A.1: 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)"	Percentages 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 3.A.2: 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 3.A.3: 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

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:

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.

3.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.

3.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 (3.3), (3.4), (3.5) and (3.6), restated here for convenience (with slight notational differences, as elaborated upon below):

$$pb_{it} = H_{it}\beta_t + X_{it}\gamma + \epsilon_{it}, \tag{3.B.1}$$

$$\epsilon_{it} = \mu_t + \rho\epsilon_{i,t-1} + u_{it}, \qquad u_{it} \sim N(0, \sigma^2_{u_i}), \tag{3.B.2}$$

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

$$\tilde{\beta}_t = \tilde{\beta}_{t-1} + \tilde{\eta}_t, \qquad \tilde{\beta}_0 = 0, \qquad \tilde{\eta}_t \sim N(0,1), \tag{3.B.4}$$

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.

(3.B.3) and (3.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^2_u are provided. This section draws from Berger et al. (2021).

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

$$y = \chi\theta + e, \quad e \sim N(0, \Sigma), \tag{3.B.5}$$

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 $\Sigma = diag\left(\sigma_u^2 \otimes \iota_T\right)$, 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

$$a_T = A_T\left(\chi'\Sigma^{-1}y + A_0^{-1}a_0\right), \tag{3.B.6}$$

$$A_T = \left(\chi'\Sigma^{-1}\chi + A_0^{-1}\right)^{-1}. \tag{3.B.7}$$

The above can then be applied to the state-space model in equations (3.B.1)-(3.B.4): First, transform the measurement equation such that its error terms are white noise. That is, insert (3.B.2) into (3.B.1) and rewrite to obtain:

$$pb_{it}^* = H_{it}^*\beta_t + X_{it}^*\gamma + u_{it}, \tag{3.B.8}$$

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 (3.B.3) into (3.B.8) yields

$$pb_{it}^* = H_{it}^*\beta_0 + H_{it}^*\sigma_\eta\tilde{\beta}_t + X_{it}^*\gamma + u_{it}, \quad u_{it} \sim N(0, \sigma_{u_i}^2), \tag{3.B.9}$$

which can be written as

$$\underbrace{pb_{it}^*}_{y_{it}} = \underbrace{\begin{bmatrix} H_{i,t}^* & H_{i,t}^*\tilde{\beta}_t & X_{it}^* \end{bmatrix}}_{\chi_{it}} \underbrace{\begin{bmatrix} \beta_0 \\ \sigma_\eta \\ \gamma \end{bmatrix}}_{\theta} + u_{it}. \tag{3.B.10}$$

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 (3.B.6) and (3.B.7).

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

$$y_t = H_t s_t + e_t, \qquad e_t \sim MN(0_N, R), \tag{3.B.11}$$

$$s_t = F s_{t-1} + K_t v_t, \qquad s_0 \sim N(b_0, V_0), \qquad v_t \sim N(0, Q), \tag{3.B.12}$$

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 (3.B.9) to obtain, together with the state equation (3.B.4), the conditional state-space model for $\tilde{\beta}_t$:

$$\overbrace{pb_{it}^* - H_{it}^*\beta_0 - X_{it}^*\gamma}^{y_{it}} = \overbrace{H_{it}^*\sigma_\eta}^{H_t} \overbrace{\tilde{\beta}_t}^{s_t} + \overbrace{u_{it}}^{e_{it}}, \qquad u_{it} \sim N(0, \overbrace{diag(\sigma_u^2)}^{R}), \tag{3.B.13}$$

$$\underbrace{\tilde{\beta}_t}_{s_t} = \underbrace{I_m}_{F} \underbrace{\tilde{\beta}_{t-1}}_{s_{t-1}} + \underbrace{I_m}_{K_t} \underbrace{\tilde{\eta}_t}_{v_t}, \qquad \tilde{\eta}_t \sim N(0, \underbrace{I_m}_{Q}), \tag{3.B.14}$$

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

$$\overbrace{\begin{bmatrix} pb^*_{1t} - H^*_{1t}\beta_0 - X^*_{1t}\gamma \\ \vdots \\ pb^*_{Nt} - H^*_{Nt}\beta_0 - X^*_{Nt}\gamma \end{bmatrix}}^{y_t} = \overbrace{\begin{bmatrix} H^*_{1t}\sigma_\eta \\ \vdots \\ H^*_{Nt}\sigma_\eta \end{bmatrix}}^{H_t} \overbrace{\begin{bmatrix} \tilde{\beta}^1_t \\ \vdots \\ \tilde{\beta}^m_t \end{bmatrix}}^{s_t} + \overbrace{\begin{bmatrix} u_{1t} \\ \vdots \\ u_{Nt} \end{bmatrix}}^{e_t}, \tag{3.B.15}$$

$$\overbrace{\begin{bmatrix} u_{1t} \\ \vdots \\ u_{Nt} \end{bmatrix}}^{e_t} \sim \left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \overbrace{\begin{bmatrix} \sigma^2_{u_1} & & \\ & \ddots & \\ & & \sigma^2_{u_N} \end{bmatrix}}^{R} \right), \tag{3.B.16}$$

$$\underbrace{\tilde{\beta}_t}_{s_t} = \underbrace{I_m}_{F} \underbrace{\tilde{\beta}_{t-1}}_{s_{t-1}} + \underbrace{I_m}_{K_t} \underbrace{\tilde{\eta}_t}_{v_t}, \tag{3.B.17}$$

$$\underbrace{\tilde{\eta}_t}_{v_t} \sim N(0, \underbrace{I_m}_{Q}). \tag{3.B.18}$$

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 (3.B.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^2_u, are drawn.

Note that, given draws of θ and β_t, ϵ_{it} and its lags are known. Thus, (3.B.2) breaks down to a conditional linear regression model, where ρ, μ and σ^2_u 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 $\sigma^2_{u_i} \sim IG(c_{0,i}, C_{0,i})$, 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, $p(\rho, \mu | \sigma^2_u, \tilde{\beta}, \theta, y, \chi) \sim N(a_{T,\{\rho,\mu\}}, A_{T,\{\rho,\mu\}})$

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 (3.B.2) for u for the respective country i, the posterior moments of the independent Normal-Inverted Gamma distribution are given by:

$$a_{T,\{\rho,\mu\}} = A_{T,\{\rho,\mu\}} \left(\chi'\Sigma^{-1}y + A^{-1}_{0,\{\rho,\mu\}}a_{0,\{\rho,\mu\}}\right) \tag{3.B.19}$$

$$A_{T,\{\rho,\mu\}} = \left(\chi'\Sigma^{-1}\chi + A^{-1}_{0,\{\rho,\mu\}}\right)^{-1} \tag{3.B.20}$$

$$c_{T,i} = c_{0,i} + (T-1)/2 \tag{3.B.21}$$

$$C_{T,i} = C_{0,i} + u_i'u_i/2 \tag{3.B.22}$$

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 $i = 1,2,...,N$.

3.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 3.B.3.

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

$$y_t = \phi_{1,t}y_{t-1} + \phi_{2,t}y_{t-2} + ... + \phi_{p,t}y_{t-p} + u_t, \qquad u_t \sim N(0,\Sigma), \tag{3.B.23}$$

$$\Phi_t = \Phi_{t-1} + e_t, \qquad e_t \sim N(0,Q), \tag{3.B.24}$$

$t = \{1,2,...,T_q\}$, 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 (3.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.

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 (3.B.23) and (3.B.24) as

$$y_t = (I_M \otimes X_t)\Phi_t + u_t, \qquad u_t \sim N(0,\Sigma), \tag{3.B.25}$$

$$\Phi_t = \Phi_{t-1} + e_t, \qquad e_t \sim N(0,Q), \tag{3.B.26}$$

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 (3.B.25) and (3.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 (3.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).

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,

$$p(Q|\Phi, \Sigma, y) \sim IW(Q_1, T_1), \tag{3.B.27}$$

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.

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

$$p(\Sigma|\Phi, Q, y) \sim IW(\Sigma_1, T_\Sigma), \tag{3.B.28}$$

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.

3.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

$$pb_{i,T+h} = \hat{\alpha}_i + H_{i,T+h}\hat{\beta}_{T+h} + X_{i,T+h}\hat{\gamma} + \epsilon_{i,T+h}, \tag{3.B.29}$$

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 3.B.1 and 3.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 the realizations of $\epsilon_{i,T+h}^r$ for each respective draw: From equation (3.B.2), it follows that $\epsilon_{i,T+h}^r = \mu_{T+h}^r + \rho^r\epsilon_{i,T+h-1}^r + u_{i,T+h}^r$. In the benchmark specification, μ_{T+h}^r is set to μ_T^r. However, a second alternative, of setting $\mu_{T+h} = 0$,

[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.

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_i^2, $i = 1, 2, ..., N$, 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):

$$debt_{i,T+h} = \frac{1 + iir_{i,T+h}}{1 + (\Delta y_{i,T+h} + \pi_{i,T+h})} + pb_{i,T+h} + sfa_{i,T+h}, \tag{3.B.30}$$

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 $t = 1, 2, ..., T$.

3.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 3.2. The FRF is given by

$$pb_{it} = \alpha_i + X_{it}\gamma + \epsilon_{it}, \quad \epsilon_{it} \sim N(0, \sigma_\epsilon^2), \tag{3.B.31}$$

$i = 1, 2, ..., N, t = 1, 2, ..., T$. 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 (3.7):

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + u_t, \qquad u_t \sim N(0, \Sigma), \tag{3.B.32}$$

$t = \{1, 2, ..., T_q\}$, 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 3.2.4, the main difference being that, unlike for the Bayesian benchmark model, parameter uncertainty is not directly incorporated in the fiscal projection exercise.

3.B.5 Published version of the paper

The published version of this article, published online on 20 February 2023 by Taylor & Francis in the journal *Applied Economics*, is available here: https://doi.org/10.1080/00036846.2023.2174500.

Chapter 4

Government spending effects on the business cycle in times of crisis

with Tino BERGER[1]

Abstract

The literature on fiscal multipliers has long established a positive impact of public spending on output. However, the size of this effect strongly depends on the employed identification strategy. Moreover, fiscal multipliers are uninformative as regards the state of the economy. Using counterfactual scenario analyses based on a conditional forecast algorithm in combination with the Beveridge-Nelson decomposition, we address both issues by assessing the effectiveness of public spending in terms of its influence on the output gap. Our approach is independent of the chosen identification strategy and allows us to make (quantitative) statements about potential downsides from public spending measures by looking at its effects on the business cycle. Using a US dataset and analyzing hypothetical government spending scenarios in times of historical crises, we find that, to avoid an overheating of the economy in combination with high inflation and public debt, the dosage of fiscal stimulus is crucial for targeted fiscal policy measures and depends on the severity of the crisis.

Keywords: Fiscal policy, output gap, conditional forecast, scenario analysis, Bayesian vector autoregression

JEL Codes: E62, E37, C53

[1]University of Göttingen.

4.1 Introduction

In times of the COVID-19 crisis, the energy crisis related to the Ukrainian-Russian war as well as long-term developments such as climate and demographic change in developed economies, government spending policies have many proponents. However, the excessive use of fiscal policy for stabilization and long-term purposes deteriorates public finances of many already debt-struck economies around the globe, potentially alarming financial markets, thereby worsening governments' refinancing conditions and thus their maneuverability to address policy goals (fiscal space). Other possible drawbacks include inflationary effects stemming from enhanced aggregate demand as well as potential unintended redistribution effects.

The fact that there is a noteworthy potential downside to excessive fiscal stimulus raises the question of public spending's effectiveness in terms of fulfilling policy objectives. These objectives center, aside from employment and (more controversial) redistribution goals, around the evolution of domestic output. Fiscal multipliers have been frequently used to assess the effectiveness of government spending in this way. However, their size strongly depends on the way the spending shocks are identified (see Ramey, 2019 and Ramey, 2016 for excellent expositions of the fiscal multiplier literature). Secondly, government spending multipliers give no indication with respect to fiscal policies' impact on the business cycle, thereby neglecting important implications for policymakers: For example, while according to data of the Federal Reserve Bank of St. Louis real GDP growth in 2020Q2 in the United States plummeted as a result of the COVID-19 outbreak, it jumped right back to +8% in the very next quarter.[2] The picture drawn here – that the economy was back on track (or beyond) in 2020Q3 – is highly misleading as it neglects the persistent character of the business cycle. Based on output gap data from the US Congressional Budget Office, the recovery was completed as late as 2021Q4.[3] The fact that the output gap is much better suited in terms of capturing the persistence in the business cycle therefore makes it the superior measure for policymaking: Central bankers might be confronted with the decision whether to raise or reduce interest rates (expand or shrink asset purchase programs). Knowing the state of the economy, in particular whether there is a positive or negative output gap, will determine whether policy rates are raised or lowered (whether balance sheets are shrunk or expanded). To identify the unobservable output gap, a stream of recent literature has argued in favor of using multivariate approaches (see e. g. Barigozzi and Luciani, 2021, Morley and Wong, 2020 or Berger et al., 2023).

In this paper, we try to overcome the two above-mentioned problems and analyze the business cycle effects of government spending policies in the United States. In particular, we employ scenario (counterfactual) analyses for some major crises in US history to assess public spending's impact on the output gap in a unified

[2] For details on the data, see appendix 4.A.

[3] The respective data can be found at the Federal Reserve Bank of St. Louis (FRED) database (here).

framework, operationalized using the multivariate Beveridge-Nelson (BN) decomposition.

By tackling the two above-mentioned problems, we contribute to the literature on the efficacy of government spending policies: First, we perform scenario analyses which are "agnostic" in the sense that they are independent of the identification strategy used to identify structural shocks (see Waggoner and Zha, 1999 or Blake and Mumtaz, 2015). This avoids that results be dependent on the identification strategy (for differences in fiscal policy efficacy arising due to differing identification assumptions in the fiscal multiplier literature, see e. g. Caldara and Kamps, 2017). Second, by imposing various fiscal policy paths in the conditional forecasting exercise and then computing implied paths of the output gap identified from a BN decomposition, we can make statements about fiscal policies' effect on the business cycle.[4]

We find that, indeed, public spending positively affects output and reduces unemployment. However, the potential downside from overspending, that is, an overheating economy with rising inflation and debt levels, implies that the dosage of fiscal stimulus matters to achieve policy goals.

The remainder of the paper is structured as follows. Section 4.2 elaborates on the multivariate BN decomposition and the scenario analyses (conditional forecasts). Section 4.3 covers the estimation strategy and the data, while in section 4.4, the empirical results are presented. Section 4.5 concludes.

4.2 Methodology

This section elaborates on the multivariate Beveridge-Nelson (BN) decomposition as well as the conditional forecast algorithm employed to compute the counterfactual scenarios.

4.2.1 The multivariate Beveridge-Nelson decomposition

In this section, we lay out the multivariate BN decomposition to compute the output gap. In particular, we follow Morley and Wong (2020) and identify the output gap as the cyclical component of the multivariate BN decomposition of the output series.

According to Beveridge and Nelson (1981), the trend of a time series y_t, t={1,2,...,T} can be defined as

$$\tau_t = \lim_{h \to \infty} \mathbb{E}_t[y_{t+h} - h\mu], \tag{4.2.1}$$

[4]Although this is not per se a *ceteris paribus* contemplation, we do get an idea of how different spending paths affect the output gap *given paths of tax revenues and monetary policy*, as elaborated upon below. Thus, grounding thoughts on fiscal effectiveness on differences in fiscal scenarios seems reasonable.

where h is the (long-run) forecast horizon and μ is a time-invariant drift. The BN cycle is then obtained as

$$c_t = y_t - \tau_t. \tag{4.2.2}$$

To compute the multivariate BN cycle as in Morley and Wong (2020), consider a standard VAR(p) model in reduced form, written as

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + u_t, \quad u_t \sim N(0, \Sigma), \tag{4.2.3}$$

$t = \{1, 2, ..., T\}$, where y_t is a $N \times 1$ vector of demeaned stationary endogenous variables, including government spending, output, inflation, the government's tax revenues, the unemployment rate and the interest rate on three-month treasury bills. All variables are transformed to stationarity (see appendix for details). ϕ_j, $j = 1, 2, ..., p$ are $N \times N$ coefficient matrices corresponding to the respective lag matrix y_{t-j} and u_t is a $N \times 1$ vector of reduced-form shocks. Following Morley and Wong (2020), express (4.2.3) in companion form:

$$\boldsymbol{Y_t} = \boldsymbol{F}\boldsymbol{Y_{t-1}} + \boldsymbol{H}u_t, \tag{4.2.4}$$

where $\boldsymbol{Y_t} \equiv \{y_t', y_{t-1}', ..., y_{t-p+1}'\}'$, $\boldsymbol{F}$ is the companion matrix, and $\boldsymbol{H}$ is a matrix mapping the reduced-form errors to the companion form. The BN trend and cycle are then given by

$$\tau_t = \boldsymbol{Y_t} + \boldsymbol{F}(\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{Y_t}, \tag{4.2.5}$$

$$c_t = -\boldsymbol{F}(\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{Y_t}. \tag{4.2.6}$$

Assuming that the output variable is the jth element of y_t in (4.2.3), the period-t output gap is the jth element of c_t.

4.2.2 Assessing fiscal policy based on scenario analysis

In this section, we lay out the methodology used to compute the counterfactuals. More precisely, given the multivariate (VAR) structure of the model, we can employ conditional multivariate forecasts (scenarios) and compute the output gap forecasts (scenarios) implied by the forecasts for the endogenous variables.

Assume that in period T, we have information about the future path of the fiscal instruments. That is, we know the $T+1, T+2, ..., T+h$ values of our government spending and tax revenue variables, where h is the number of periods we have information on the fiscal instruments for. That is, h is the scenario horizon. Now suppose we are interested in knowing the values in $T+1, T+2, ..., T+h$ for the model's other variables. Noting that, technically, such a scenario analysis is simply a conditional forecast where we employ information on the fiscal instruments to forecast the remaining variables, our approach, outlined below, draws heavily from Waggoner and

Zha (1999), Blake and Mumtaz (2015), Higgins et al. (2016) and Berger et al. (2023). First, rewrite equation (4.2.3) as

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + A_0^{-1}\epsilon_t, \quad \epsilon_t \sim N(0, I), \tag{4.2.7}$$

having employed the relation $u_t = A_0^{-1}\epsilon_t$, where A_0^{-1} is a structural impact multiplier matrix, with $\Sigma = A_0^{-1}\left(A_0^{-1}\right)'$. That is, A_0^{-1} is the lower-triangular Cholesky factor of Σ. ϵ_t are structural shocks.[5]

Rewriting (4.2.7) in companion form yields

$$\boldsymbol{Y}_t = \boldsymbol{F}\boldsymbol{Y}_{t-1} + \boldsymbol{H}\boldsymbol{A_0}^{-1}\epsilon_t, \tag{4.2.8}$$

where $\boldsymbol{H}$ is a matrix mapping the reduced-form shocks $\boldsymbol{A_0}^{-1}\epsilon_t$ to the companion form, that is,

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{I_N} \\ \boldsymbol{0}_{N(p-1)\times N} \end{bmatrix}.$$

Iterating (4.2.8) h steps forward yields

$$\boldsymbol{Y_{t+h}} = \boldsymbol{F}^h\boldsymbol{Y}_t + \sum_{j=1}^{h}\boldsymbol{F}^{h-j}\boldsymbol{H}\boldsymbol{A_0}^{-1}\epsilon_{t+j}. \tag{4.2.9}$$

Letting $\boldsymbol{F}^k\boldsymbol{H}\boldsymbol{A_0}^{-1} := \boldsymbol{M}_k$ be the impulse response matrix at horizon k, this can be written as

$$\boldsymbol{Y}_{t+h} = \boldsymbol{F}^h\boldsymbol{Y}_t + \sum_{j=1}^{h}\boldsymbol{M}_{h-j}\epsilon_{t+j}, \tag{4.2.10}$$

$$\sum_{j=1}^{h}\boldsymbol{M}_{h-j}\epsilon_{t+j} = \boldsymbol{Y}_{t+h} - \boldsymbol{F}^h\boldsymbol{Y}_t. \tag{4.2.11}$$

That is, given the model's parameters, restricting some variables in Y_{t+h} implies restrictions on the structural shocks ϵ_t. Conditional on this information as well as the estimated reduced-form parameters, collected in $\boldsymbol{F}$, the paths of the unconditioned variables can be obtained. To see this, follow Higgins et al. (2016) and assume we have an idea about the future path of the endogenous variables y and therefore restrict their future values to $\boldsymbol{Y}^* = \left(\boldsymbol{Y}_{T+1}^{*'}, ..., \boldsymbol{Y}_{T+h}^{*'}\right)'$, where the dimension of each $\boldsymbol{Y}_t^*$

[5] Note that the conditional forecasts are not affected by the choice of identification schemes to recover exogenous shocks (see Waggoner and Zha, 1999, Blake and Mumtaz, 2015). Thus, differences in the effectiveness of fiscal policy paths do not stem from the choice of the structural impact multiplier matrix, such that we do not need to engage in a discussion on how the chosen identification strategy affects our results. This is a major advantage compared to approaches identifying fiscal policy's effectiveness from structural VAR or DSGE models.

is $Np \times 1$ (corresponding to the companion form vector $\boldsymbol{Y}_t$). Next, define the corresponding vector of *unconditional* forecasts as $\boldsymbol{Y}^u = \left(\boldsymbol{Y}^{u'}_{T+1}, ..., \boldsymbol{Y}^{u'}_{T+h}\right)'$. Given these definitions, define $r := \boldsymbol{Y}^* - \boldsymbol{Y}^u$, $\epsilon := (\epsilon'_{T+1}, \epsilon'_{T+2}, ..., \epsilon'_{T+h})'$ and

$$\boldsymbol{R} := \begin{bmatrix} \boldsymbol{M_0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{M_1} & \boldsymbol{M_0} & \cdots & \boldsymbol{0} \\ \vdots & & & \\ \boldsymbol{M_{h-1}} & \boldsymbol{M_{h-2}} & \cdots & \boldsymbol{M_0} \end{bmatrix}.$$

Note that we can conveniently summarize the restrictions imposed on the future values of the endogenous variables y, conditional on the reduced-form parameters of the VAR. To do this, collect the reduced-form parameters in $a := (vec(\phi_1)', vec(\phi_2)', ..., vec(\phi_p)', vec(A_0^{-1}))'$ and stack (4.2.11) over the whole forecast horizon $\{1, 2, ..., h\}$. That is, express (4.2.11) as

$$R(a)\epsilon = r(a). \tag{4.2.12}$$

In general, one is interested in restricting only a subset of the endogenous variables. Similar to Higgins et al. (2016) and Blake and Mumtaz (2015), define $\tilde{R}$ and $\tilde{r}$ as the respective matrices where the rows corresponding to the unrestricted variables are excluded and define q as the number of endogenous variables with restricted future paths. We can then rewrite (4.2.12) as

$$\tilde{R}(a)\epsilon = \tilde{r}(a), \tag{4.2.13}$$

where $\tilde{R}(a)$ is of dimension $qph \times Nh$, $\tilde{r}(a)$ is $qph \times 1$, and $qh \leq Nh$. As shown in Doan et al. (1984), the structural shocks ϵ can be estimated using ordinary least squares. That is,

$$\hat{\epsilon} = \tilde{R}'(\tilde{R}\tilde{R}')^{-1}\tilde{r}. \tag{4.2.14}$$

With an estimate of the structural shocks $\hat{\epsilon}$ at hand, the conditional forecasts can easily be recovered using equation (4.2.8).

Since we are interested in the effectiveness of public spending policies, we assume various paths of government spending and analyze the implied scenarios for the remaining variables. However, we constrain the path of tax revenue growth so as to exclude the possibility that the scenario results are driven by changes on the revenue side. At the same time, we condition the monetary policy variable in the model – the first differenced three-month treasury bill rate – such that differing scenario forecasts neither result from changes in monetary policy. Thus, the unconditioned variables are output growth, inflation and the unemployment rate. Details on the conditions imposed in each of the scenarios are outlined in section 4.4.2.

With the scenario paths of all variables at hand, we can compute our estimate of the output gap in periods $T+1, T+2, ..., T+h$ as in Berger et al. (2023):

$$c_{t+1} = -\boldsymbol{F}(\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{Y_{t+1}}. \tag{4.2.15}$$

Again, assuming output growth is the jth variable in y_t means that the output gap is the jth element in c_{t+1}.

4.3 Empirical framework

In this section, the estimation procedure as well as the data are outlined.

4.3.1 BVAR estimation using a Minnesota dummy observation prior

This section lays out the estimation approach for the VAR(p) model. In particular, we use Bayesian estimation techniques, drawing from Berger et al. (2023) and Morley and Wong (2020).

Consider again the standard VAR(p) model in reduced form of equation (4.2.3), restated here for convenience:

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

$t = \{1, 2, ..., T\}$. With each ϕ_j, $j = 1, 2, ..., p$ being of dimension $N \times N$, the high number of parameters to be estimated for increasing N and p motivates Bayesian estimation techniques. Following Morley and Wong (2020), a natural conjugate Minnesota dummy observation prior is employed, which applies shrinkage on the model's parameters and implies that the posterior means are obtainable as closed-form solution. For illustrative purposes, consider (4.2.3) in expanded form:

$$y_t = \begin{bmatrix} \phi_1^{1,1} & \cdots & \phi_1^{1,N} & \phi_2^{1,1} & \cdots & \phi_2^{1,N} & \cdots & \cdots & \phi_p^{1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ \phi_1^{N,1} & \cdots & \phi_1^{N,N} & \phi_2^{N,1} & \cdots & \phi_2^{N,N} & \cdots & \cdots & \phi_p^{N,N} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ \vdots \\ u_{N,t} \end{bmatrix} \tag{4.3.1}$$

As outlined in Morley and Wong (2020), the first two prior moments of $\phi_i^{j,k}$, the slope coefficient corresponding to the i^{th} lag of the k^{th} variable in the j^{th} equation, are set as

$$\mathbb{E}\left[\phi_i^{j,k}\right] = 0, \tag{4.3.2}$$

$$Var\left[\phi_i^{j,k}\right] = \begin{cases} \dfrac{\lambda^2}{i^2}, & j = k, \\ \dfrac{\lambda^2}{i^2}\dfrac{\sigma_j^2}{\sigma_k^2}, & \text{else.} \end{cases} \tag{4.3.3}$$

Thus, we follow Bańbura et al. (2010) in setting the first prior moment to 0 for all slope parameters due to their "substantial mean reversion", following from the imposition of stationarity on all series. The coefficient λ, which according to Bańbura et al. (2010) governs the prior's "overall shrinkage", is obtained from minimizing the corresponding pseudo-out-of-sample one-period-ahead forecast errors of the real GDP growth series, as in Morley and Wong (2020). Clearly, our loss function centers around real GDP as we want to optimize the model's forecasting capacity with respect to the output gap.

The choice of an out-of-sample loss function is to avoid overfitting that might be more likely to occur when minimizing an in-sample forecast error (once again, see Morley and Wong, 2020). The one-step-ahead root mean squared forecast error is computed recursively, with an initial sample covering the first 80 observations (that is, the first 20 years of the sample), then adding one observation in turn up to period $T - 1$.

Intuitively, λ approaching zero is equivalent to the assumption that the variables tend to be independent white noise processes. Further note that a common feature of the Minnesota prior is the $\frac{1}{i^2}$ term in the prior variances of $\phi_i^{j,k}$, which implies that longer lags are shrunk more towards the mean, that is, towards zero. The σ_j^2 and σ_k^2 terms stem from AR(4) processes of the respective variables, estimated with ordinary least squares, as is common in the literature (see, among others, Berger et al., 2023 and Bańbura et al., 2010). Lastly, note that working with demeaned variables is equivalent to employing constants with a flat prior in each equation.

Following Morley and Wong (2020) and Del Negro and Schorfheide (2011), the model can be estimated by first embedding the above-specified prior by adding dummy observations to the data set and then simply running least squares on the extended data set, which is feasible due to the natural conjugacy of the prior.

4.3.2 Data

As mentioned above, our BVAR model includes six variables, namely real GDP, CPI inflation, the unemployment rate, the three-month treasury bill rate, real government current receipts and real government spending, motivated by standard choices in the fiscal multiplier literature (see for example Caldara and Kamps, 2017 or Ramey, 2019). Our quarterly dataset covers observations from 1952Q1 to 2022Q2.

Non-stationary series are transformed to stationarity. Sources and transformations of all series are provided in appendix 4.A.

4.4 Empirical results

In this section, our empirical results are presented. In section 4.4.1, we show our *ex-post* output gap, based on the full sample information. Section 4.4.2 describes the considered scenarios and lays out the results of the counterfactual analysis.

4.4.1 Ex-post output gap results

This section outlines the results of the estimated output gap based on the full information dataset. That is, we use all observations, from 1952Q1 to 2022Q2, and apply equation (4.2.6) to get the full information estimate of the Beveridge-Nelson output gap.

Figure 4.1 shows this result. The thick blue line represents the posterior mean of the output gap estimate, with the blue shaded area being the 90% credible set, and the gray shaded vertical areas representing NBER recession dates. As can be see from this figure, our estimated output gap captures the NBER recessions quite well, with the First Oil Crisis, the recession of 1981 to 1982, the Great Recession and the COVID-19 Recession being the most severe ones according to the respective business cycle troughs. Moreover, note that our output gap estimate is quite similar to both those in Berger et al. (2023) and Morley and Wong (2020), who use a higher number of variables and, in the case of Berger et al. (2023), a higher data frequency for most variables.

4.4.2 The fiscal scenarios

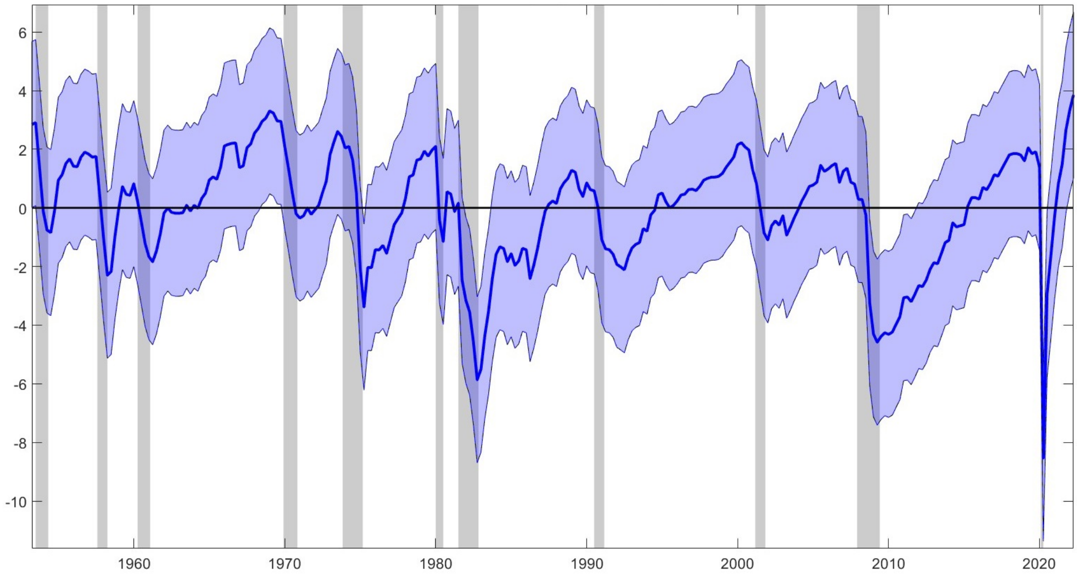
In this section, we analyze hypothetical macroeconomic consequences of various fiscal scenarios for different economic crises in US history. That is, unlike in section 4.4.1, we now look into counterfactuals that – according to our model – would have occurred for different paths of government spending. As explained above, in each of the scenarios, we control for the path of tax revenue and our model's nominal interest rate measure, both of which are set to follow their actual (ex-post observed) paths in each of the analyzed crisis scenarios.

Section 4.4.2 presents some details on the analyzed fiscal scenarios – from restrictive to "super-expansive". Section 4.4.2 briefly elaborates on the chosen crises. Finally, the scenario results are presented in section 4.4.2.

The scenarios

We start by outlining the scenarios analyzed for each of the economic crises. For each crisis, we present four scenarios, which all differ according to the "expansiveness" of the government spending path. For all scenarios, using the ex-post information on

FIGURE 4.1: Output gap based on full sample information



Notes: The thick blue line represents the posterior mean of the estimated output gap based on the full information set, with the surrounding shaded area being the 90% credible interval of that estimate. The gray shaded areas indicate NBER recession dates.

the reduced-form parameters (that is, using all information up to the final observation in 2022Q2), we consider the hypothetical case of forecasting output growth, the unemployment rate, consumer price inflation and, by implication, the output gap, *at the time of the respective crisis*, conditional on the assumption that our government spending variable takes a certain path. At the same time, we control for the paths of tax revenues and the nominal interest rate, which we restrict to have the same values in each scenario. Thus, differences in the paths of the unrestricted variables – and particularly the output gap – will not result from different tax or monetary policies. The four scenarios can be described in the following way:

- **Actual spending path:** In this scenario, government spending follows its actual ex-post path. Assume we are interested in knowing what our model would predict for the business cycle at the height of the Global Financial Crisis, say at the end of 2008. The *actual spending path* scenario answers the question: According to our model, what values would the output gap take in 2009-2013 *if we knew the path government spending would take in 2009-2013* (and given the ex-post reduced-form parameter estimates and the paths of tax revenues and the nominal interest rate)?

- **Restrictive spending path:** In this scenario, we assume that government spending growth is 1 percentage point lower than it actually was during the whole scenario (forecast) horizon. Our Global Financial Crisis question becomes: How would the output gap be affected if in 2009-2013, government spending growth would be 1 percentage point lower than it actually was (given the same ex-post reduced-form parameter estimates and the paths of tax revenues and the nominal interest rate as in the actual spending path scenario)?

- **Expansive spending path:** This time, we assume government spending growth is 1 percentage point higher than it actually was for the whole scenario (forecast) horizon.

- **Super-expansive spending path:** Finally, we consider a "super-expansive" public spending scenario, where government spending growth is 5 percentage points higher than it actually was during the whole scenario (forecast) horizon.

For all scenarios, tax revenue growth and the first difference of the three-month treasury bill rate – the transformations we use in the model – will be conditioned to follow their actual ex-post paths.

The crises

The previously outlined scenarios are analyzed for a variety of crises in US history. In particular, we look into the four most severe crises as defined by our estimates of their business cycle troughs, displayed in figure 4.1. For each of the crises, we will briefly elaborate on the policy measures in place at the time as well as on the dating of the crises as defined by the NBER.[6] Finally, note that we assume some sluggishness in the implementation of fiscal measures, thus starting our scenario analyses two quarters after the respective recession start date. The scenario start dates are included in the following recession summaries.

- **The First Oil Crisis:** At the time of the First Oil Crisis, economic policy was dominated by the Federal Reserve, which was particularly concerned with countering the pronounced inflation dynamics: The monetary tightening certainly did not support the recovery. On the fiscal side, policymakers finally used tax cuts to stabilize the economy (see Blinder, 2022). According to the NBER, the First Oil Crisis lasted from November 1973 to March 1975. Following the logic described above, we start our scenario analysis in the second quarter of 1974.

- **The 1981-1982 recession:** When the 1981-1982 recession hit, the Federal Reserve was once again dominating economic policy, and once again mainly concerned with bringing down high inflation levels. However, President Reagan brought fiscal policy back to the center of attention: To counter the deep recession, the Reagan government implemented enormous tax cuts in 1981, 1982 and 1983, amounting, according to Blinder (2022), to a 23% personal income tax rate reduction in total, thus strongly weighing on the government's budget

[6]See here for the business cycle dating. For the description of the economic policies in place during the recessions, we borrow from Blinder (2022), who provides an excellent overview on historical monetary and fiscal policies in the US.

balance. Government spending still played no prominent role in terms of stabilization policies and was even reduced to partly finance the arising budget deficit. However, due to increased military spending expenses that incurred at the same time, this deficit reduction was negligible, implying soaring debt-to-GDP levels (see Blinder, 2022). According to the NBER, the 1981-1982 recession lasted from July 1981 to November 1982, implying a start date for our counterfactuals in 1982Q1.

- **The Great Recession:** Roughly 25 years later, monetary policy still held supremacy with respect to the conduct of stabilization policy: The primary response to the Great Recession was a massive reduction in the policy rate of more than 5 percentage points, combined with other measures such as *quantitative easing* and *forward guidance*. On the fiscal side, things changed once President Obama took office, who quickly implemented the *American Reinvestment and Recovery Act* (ARRA) – a massive fiscal stimulus package of approximately 5% of GDP, which was a combination of expansionary spending and tax measures. As laid out in Blinder (2022), the ARRA was far from uncontroversial: While some prominent voices, among them Paul Krugman and Christina Romer, argued that the stimulus program was not sufficient given the size of the recession, the Obama administration faced a lot of headwind particularly from the Republican side, whose criticism focused on the spending components of the stimulus package. After the midterms, with the Republicans having reclaimed the House of Representatives, fiscal policy even became contractionary towards the beginning of the 2010s, much to the disliking of the incumbent chairman of the Federal Reserve, Ben Bernanke, who was suggesting the Fed and government to move in lock step to further soften the recessionary blow (see Blinder, 2022). With NBER recession start and end dates in December 2007 and June 2009, we start our Great Recession scenarios in 2008Q2.

- **The COVID-19 Recession:** The policy response to the COVID-19 Recession was different again, as with the major monetary expansion in response to the Great Recession, the Fed's policy rate was close to the zero lower bound. The Fed used what maneuverability it still had and reduced rates further, simultaneously once again resorting to quantitative easing and forward guidance. The monetary dominance of former crises was over, though: Most prominently, the incumbent governments passed a variety of fiscal stimulus packages, among them the *Coronavirus Aid, Relief, and Economic Security Act* and later on the *American Relief Plan*, with a total of approximately 6 trillion USD, together more than 27% of 2021 GDP. Unlike in the Great Recession, the biggest components of these stimulus packages were tax cuts and transfer payments (for more details, see Blinder, 2022). According to the NBER, the COVID-19

Recession lasted from February to April 2020, implying a scenario start date in 2020Q3.

In conclusion, government spending was not the main contributor to the stabilization policies for the crises at hand, even though things have shifted more towards fiscal policy measures (both on the revenue and expenditure front) since the Great Recession. With the following scenario analyses, we address the question whether different government spending policies would have led to different economic outcomes, especially in terms of the speed of recovery.

Scenario results

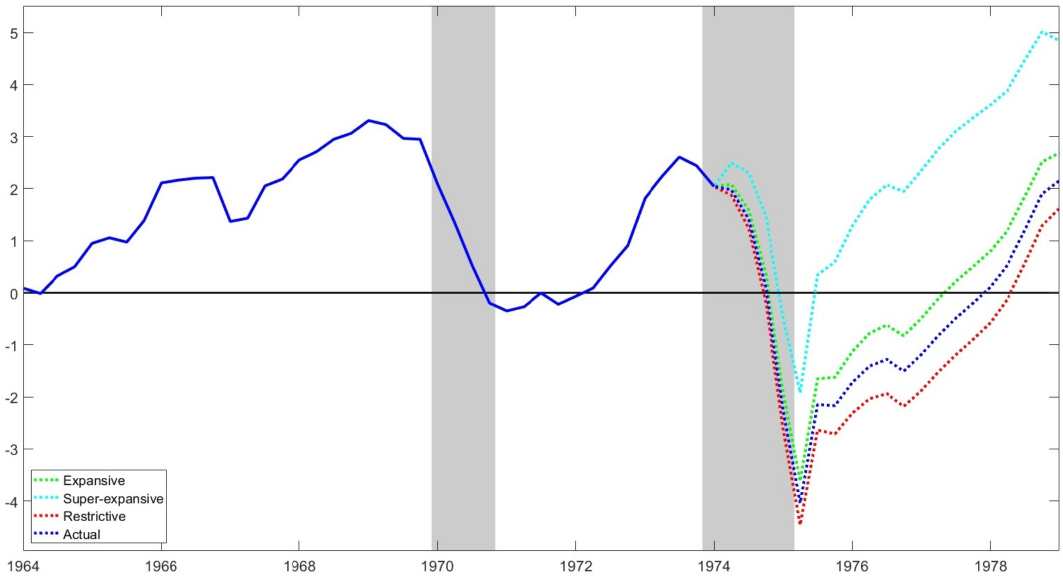
This section presents the scenario results for all of the above-mentioned crises. For all but the COVID-19 crisis, the scenario (conditional forecast) horizon is 5 years (20 quarters). Since the conditions for the fiscal and monetary variables are based on their actual ex-post values, the COVID-19 scenario horizon is restricted to 8 quarters, that is, the scenario ends in 2022Q2, which is the last observation in the sample.

Figure 4.2 displays the scenario results for the First Oil Crisis. The thick blue line represents the mean path of the ex-post output gap (as presented in figure 4.1). The scenarios start in 1974Q2, that is, early in the First Oil Crisis, as defined by the NBER. The dashed blue line indicates the scenario where government spending follows its true (ex-post) path ("actual spending path" scenario, see section 4.4.2). The red, green and cyan dashed lines represent the restrictive, expansive and super-expansive scenarios, respectively. As can be seen, had the US government done more to counter the crisis (by means of increased government spending) early on, the recession would have been less pronounced. For example, in the case of the super-expansive scenario, the business cycle trough would have been at approximately -1.9% instead of -3.4% in the case of the full information estimate (displayed in figure 4.1), while the expansion phase would have been reached as early as 1975Q3 instead of 1978Q1.

However, this highly expansionary spending path is not costless: Had the US government in fact raised public spending growth by an amount of 5 percentage points above its actual path, the public debt-to-GDP ratio (debt ratio) would have roughly risen by 21 percentage points over the analyzed five year period.[7] Moreover, with the output gap closed so quickly in the super-expansive scenario, there clearly is a danger of overheating in this case, with the output gap rising above 5% towards the end of the scenario horizon. This development is also confirmed by the implied rising inflation levels, see figure 4.B.1 in the appendix.

[7]Instead of 32% at the end of the scenario horizon (in 1978Q4), the debt ratio would have risen to 53%. However, this is just a very rough guess, assuming the spending-induced increase in aggregate demand has no feedback effects on the debt ratio. It serves – in a simplistic way – the purpose of illustrating the looming danger of overspending in terms of fiscal sustainability.

FIGURE 4.2: Output gap paths for First Oil Crisis scenario



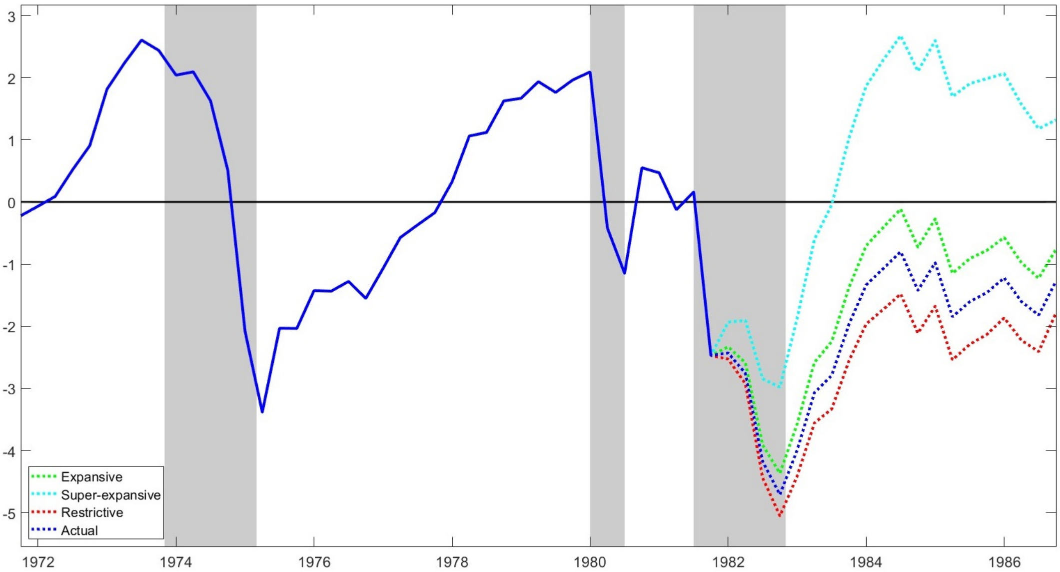
Notes: Displayed are the implied paths of the output gap resulting from four different fiscal scenarios: Actual (blue), restrictive (red), expansive (green) and super-expansive (cyan). The thick blue line indicates sample observations, dashed lines indicate scenarios paths. NBER recession dates are depicted as gray shaded areas. The scenario (forecast) horizon is 20 quarters.

A less severe increase in the output gap and therefore inflationary pressure would have occurred for a less pronounced fiscal expansion: In the expansive scenario, where each period's public spending growth is just 1 percentage point higher than its ex-post value (instead of 5 percentage points in the super-expansive scenario) the output gap towards the scenario end approaches 3% instead of 5% in the super-expansive case. This is again confirmed by lower inflation tendencies as indicated in figure 4.B.1, with quarter-on-quarter inflation roughly at 1.5% towards scenario end, instead of almost 2% in the super-expansive case. On the other hand, the lower amount of public spending means that the recession is much more severe, as indicated by the trough of the green dashed line. In fact, the recession is hardly cushioned at all, and the output gap is closed only in 1977Q3 (only two quarters before the same occurs in the actual spending scenario and according to the full information estimate of figure 4.1). At the same time, the debt ratio lies 4 percentage points above its ex-post value at scenario end.

Clearly, resorting to stabilization policies leads to the well-known trade-off between the speed of recovery and the possibility of overheating with inflationary tendencies, combined with a potential strain on fiscal solvency as a consequence of increasing debt levels. Scenario analyses of the kind presented here provide policymakers with a tool to quantitatively investigate this trade-off.

The importance of the "dosage" of fiscal stimulus, injected into the system, is apparent for the 1981-1982 recession as well. Figure 4.3 presents the output gap scenarios for this crisis. Again, the super-expansive scenario shows a much less severe downturn, with a trough around -3% (as opposed to -4% to -6% for the alternative

FIGURE 4.3: Output gap paths for 1981-1982 recession scenario



Notes: Displayed are the implied paths of the output gap resulting from four different fiscal scenarios: Actual (blue), restrictive (red), expansive (green) and super-expansive (cyan). The thick blue line indicates sample observations, dashed lines indicate scenarios paths. NBER recession dates are depicted as gray shaded areas. The scenario (forecast) horizon is 20 quarters.

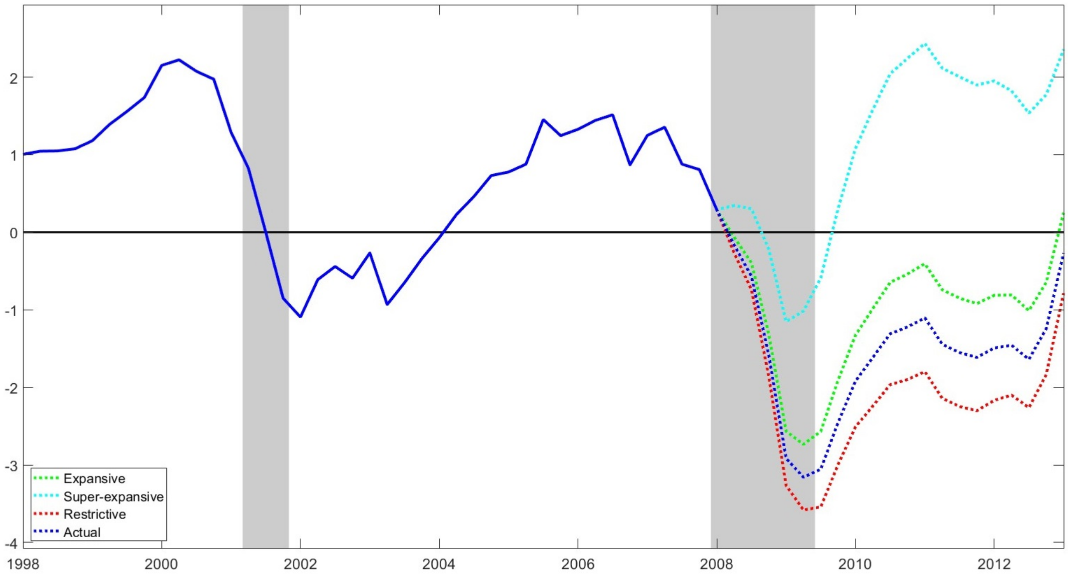
cases). This time, though, in all but the super-expansive scenario, the output gap is never closed for the duration of the scenario. This makes a stronger case for excessive fiscal spending and might be more of a justification of the (hypothetically) resulting pronounced increase in public debt.[8]

Similar findings are given for the Great Recession case in figure 4.4: A massive and persistent increase in fiscal stimulus as in the super-expansive scenario would have led to a distinctly faster recovery, both compared to the alternative scenarios as well as the full information output gap: In the super-expansive case, the recovery phase would have set in at the end of 2009 already, while – according to our model – in the actual spending scenario and for the full information case a recovery was not even in place towards the end of the scenario horizon. Again, despite a hypothetical increase in the debt ratio from 99% to 117% at the end of the scenario horizon (ignoring potential mitigating effects from higher GDP growth), there is a case for pronounced fiscal stimulus, and even more so than already occurred in response to the Great Recession. Thus, our findings are somewhat in line with the position of proponents of more fiscal stimulus, mentioned above. The case for more public spending is also confirmed by an only moderate increase in inflation (see figure 4.B.3).

In the COVID-19 case, presented in figure 4.5, our model predicts that additional positive effects on the output gap are expensively bought: Although massive public

[8]However, it should be noted that the increase in the debt-to-GDP ratio in the Reagan years was quite pronounced even without the high public spending growth rates of the super-expansive scenario. One could argue, nevertheless, that a faster recovery might at least partially offset the increase in the debt ratio due to automatic stabilizers and a higher denominator.

FIGURE 4.4: Output gap paths for Great Recession scenario



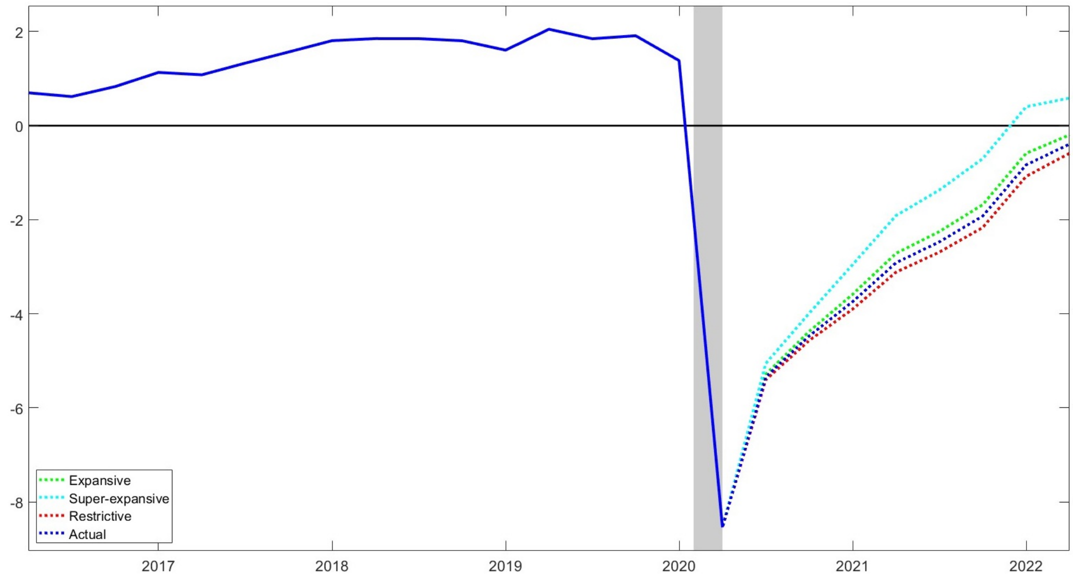
Notes: Displayed are the implied paths of the output gap resulting from four different fiscal scenarios: Actual (blue), restrictive (red), expansive (green) and super-expansive (cyan). The thick blue line indicates sample observations, dashed lines indicate scenarios paths. NBER recession dates are depicted as gray shaded areas. The scenario (forecast) horizon is 20 quarters.

spending does help to close the output gap more quickly, the difference appears almost negligible, with a difference in means against the actual spending scenario over the scenario horizon of 0.54 percentage points. Given the implied increase in the debt ratio of roughly 8 percentage points, this seems costly. However, note that the COVID-19 scenario is restricted to 8 quarters only, as mentioned above. Thus, judging the full effect of a fiscal stimulus as shown for the other crises needs to be assessed in the future. It should still be noted that this finding appears somewhat counterintuitive, given the literature on more effective fiscal multipliers at the zero lower bound (see e. g. Ramey, 2019).[9]

Finally, note that according to our model, massive fiscal stimulus has a pronounced effect on the unemployment rate, too: For example, in the case of the 1981-1982 recession, the super-expansive scenario predicts an unemployment rate of roughly 3.9% as opposed to 6.7% for the actual spending scenario and a true (ex-post) value of 6.8%. The reduction in the unemployment rate is similar for the First Oil Crisis and the Great Recession cases, again making more of a case for (even more) pronounced fiscal stimulus. As for the output gap, the unemployment effect in the COVID-19 case appears to be smaller, at least upon observation of the shorter scenario of only 2 years.

[9]However, note that while for the first three crises considered, the conditional forecasts (scenarios) are quite close to the estimated ex-post gaps, this is not the case for the COVID-19 case. Here, the recovery according to the full information gap estimate set in much faster than predicted by our model. Still, the finding that a massive increase in fiscal stimulus on the spending side does not seem to have much effect here is insightful.

FIGURE 4.5: Output gap paths for COVID-19 scenario



Notes: Displayed are the implied paths of the output gap resulting from four different fiscal scenarios: Actual (blue), restrictive (red), expansive (green) and super-expansive (cyan). The thick blue line indicates sample observations, dashed lines indicate scenarios paths. NBER recession dates are depicted as gray shaded areas. The scenario (forecast) horizon is restricted to 8 quarters since the conditions on fiscal and monetary variables are proportional to the variables' ex-post paths, which in this case are available only for 8 "future" (pseudo-out-of-sample) observations.

4.5 Conclusion

Challenges justifying a role for fiscal policy and public spending are omnipresent: Be it distortions resulting from recent crises such as the COVID-19 pandemic or the Ukrainian-Russian war, or long-term challenges such as increased financing needs in aging societies or investment requirements to tackle and adapt to climate change.

At the same time, these very challenges strongly weigh on fiscal sustainability, especially given the already pronounced debt levels in advanced economies. Lately, rising interest rates and the consequently increasing refinancing costs amplify the severity of these dynamics. Therefore, placing the trade-off between stimulating effects from public spending on the one hand and overheating as well as potentially rising debt levels on the other at the center of attention is essential for adequate policymaking.

In accordance with the literature on fiscal multipliers, we show that an increase in public spending positively affects output and reduces unemployment in times of crisis. In addition, we provide empirical evidence for a positive impact of public spending on the output gap, thereby extending the debate on fiscal efficacy from the mere discussion of output levels and growth rates to the more policy-relevant question of business cycle effects. Thus, next to the upside of fiscal expansions, our model gauges potential downsides from "overspending" in terms of an overheating economy and resulting inflationary effects, which should be evaluated on a case-by-case basis. In this light, our model hands the prudent fiscal policymaker a tool to assess the dosage of public spending measures.

Bibliography

Barigozzi, M. and Luciani, M. (2021). Measuring the output gap using large datasets. *The Review of Economics and Statistics*, pages 1–45.

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92.

Berger, T., Morley, J., and Wong, B. (2023). Nowcasting the output gap. *Journal of Econometrics*, 232(1):18–34.

Beveridge, S. and Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics*, 7(2):151–174.

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

Blinder, A. S. (2022). *A monetary and fiscal history of the United States, 1961–2021*. Princeton University Press, Princeton.

Caldara, D. and Kamps, C. (2017). The analytics of SVARs: A unified framework to measure fiscal multipliers. *The Review of Economic Studies*, 84(3):1015–1040.

Del Negro, M. and Schorfheide, F. (2011). Bayesian macroeconometrics, chapter 7 in the handbook of bayesian econometrics, ed. by jf geweke, g. koop, and hk van dijk.

Doan, T., Litterman, R., and Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, 3(1):1–100.

Higgins, P., Zha, T., and Zhong, W. (2016). Forecasting china's economic growth and inflation. *China Economic Review*, 41:46 – 61.

Morley, J. and Wong, B. (2020). Estimating and accounting for the output gap with large bayesian vector autoregressions. *Journal of Applied Econometrics*, 35(1):1–18.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. *Handbook of macroeconomics*, 2:71–162.

Ramey, V. A. (2019). Ten years after the financial crisis: What have we learned from the renaissance in fiscal research? *Journal of Economic Perspectives*, 33(2):89–114.

Waggoner, D. F. and Zha, T. (1999). Conditional forecasts in dynamic multivariate models. *The Review of Economics and Statistics*, 81(4):639–651.

Appendix

4.A Data appendix

Table 4.A.1 displays the data series employed in the BVAR. All series are taken from the Federal Reserve Bank of St. Louis database (FRED) and are transformed to stationarity. The respective transformation of the series is displayed in the table.[10]

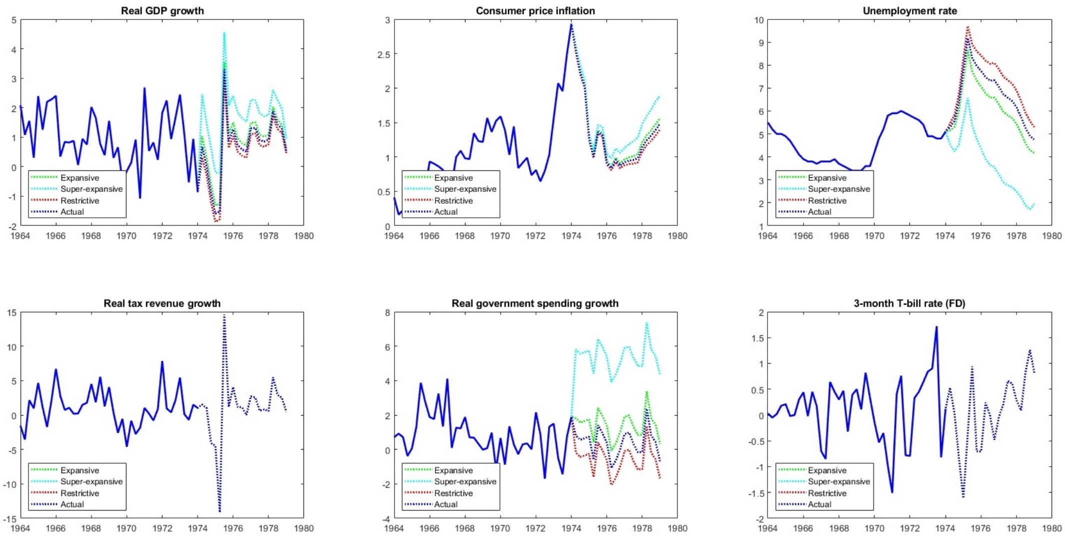
TABLE 4.A.1: Data description for the vector autoregression

Series name	Sources	Transformation
Real government spending	"Government Consumption Expenditures and Gross Investment" (FRED code: GCE), deflated by "Gross Domestic Product: Implicit Price Deflator" (FRED code: GDPDEF)	Δln
Real gross domestic product	"Gross domestic product" (FRED code: GDP), deflated by "Gross Domestic Product: Implicit Price Deflator" (FRED code: GDPDEF)	Δln
Consumer price index	"Consumer Price Index for All Urban Consumers: All Items in U.S. City Average" (FRED code: CPIAUCSL)	Δln
Real government current receipts	"Federal Government Current Receipts" (FRED code: FGRECPT), deflated by "Gross Domestic Product: Implicit Price Deflator" (FRED code: GDPDEF)	Δln
Unemployment rate	"Unemployment Rate" (FRED code: UNRATE)	$-$
Nominal interest rate	"3-Month Treasury Bill Secondary Market Rate, Discount Basis" (FRED code: TB3MS)	Δ

[10]Note that growth rates are expressed as quarter-on-quarter percentage changes.

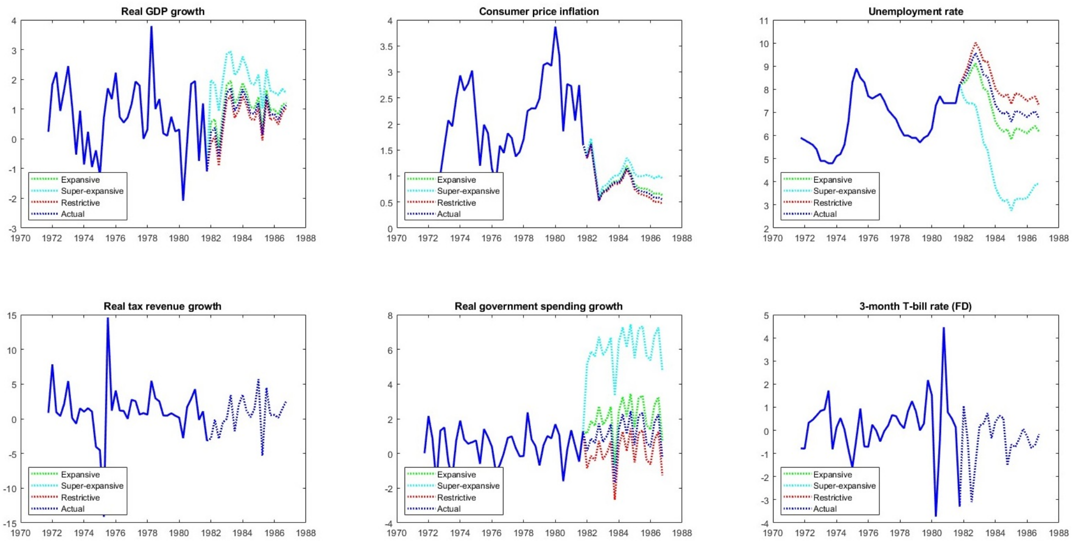
4.B Scenario plots of endogenous variables

FIGURE 4.B.1: Endogenous variable paths for First Oil Crisis scenario



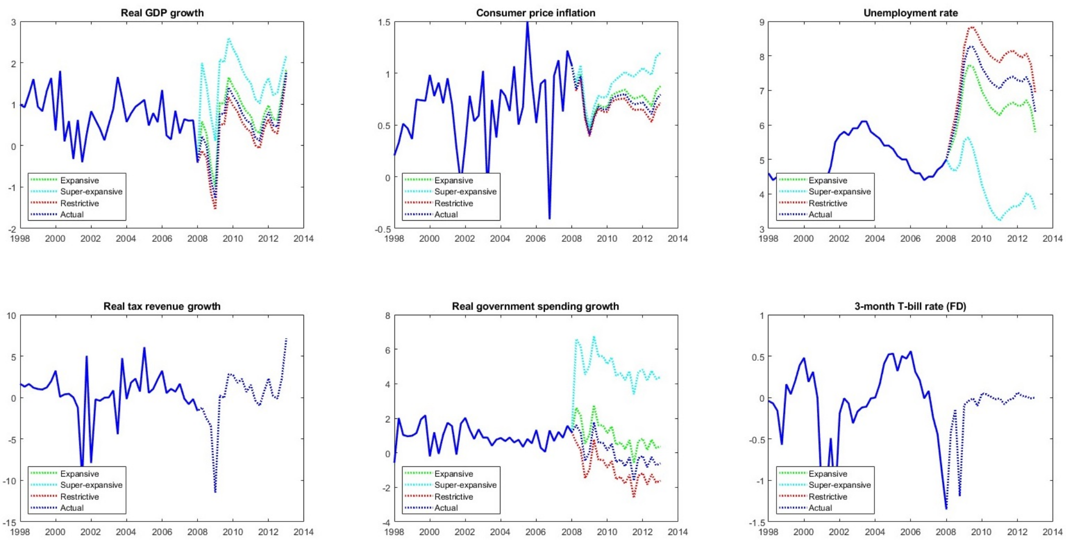
Notes: Displayed are the implied paths resulting from four different fiscal scenarios: Actual (blue), restrictive (red), expansive (green) and super-expansive (cyan). The thick blue line indicates sample observations, dashed lines indicate scenario paths.

FIGURE 4.B.2: Endogenous variable paths for 1981-1982 recession scenario



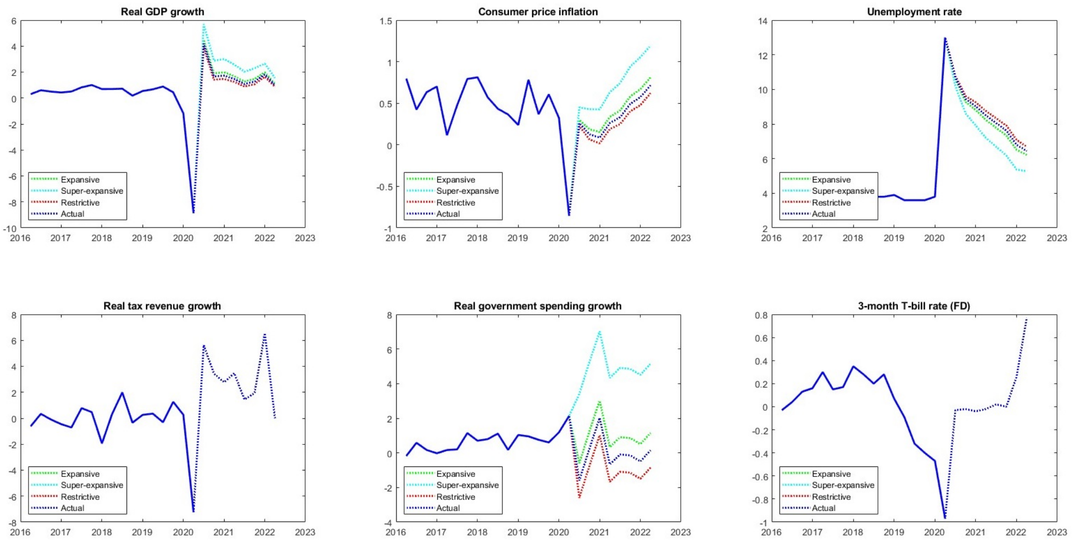
Notes: Displayed are the implied paths resulting from four different fiscal scenarios: Actual (blue), restrictive (red), expansive (green) and super-expansive (cyan). The thick blue line indicates sample observations, dashed lines indicate scenario paths.

FIGURE 4.B.3: Endogenous variable paths for Great Recession scenario



Notes: Displayed are the implied paths resulting from four different fiscal scenarios: Actual (blue), restrictive (red), expansive (green) and super-expansive (cyan). The thick blue line indicates sample observations, dashed lines indicate scenario paths.

FIGURE 4.B.4: Endogenous variable paths for COVID-19 scenario



Notes: Displayed are the implied paths resulting from four different fiscal scenarios: Actual (blue), restrictive (red), expansive (green) and super-expansive (cyan). The thick blue line indicates sample observations, dashed lines indicate scenario paths.

Appendix A

Assurance according to §12 PStO

Ph.D. program in Economics

Declaration for admission to the doctoral examination

I confirm

1. that the dissertation that I submitted

Designing fiscal policy - On the trade-off between macroeconomic stabilization and government debt sustainability

was produced independently without assistance from external parties, and not contrary to high scientific standards and integrity,

2. that I have adhered to the examination regulations, including upholding a high degree of scientific integrity, which includes the strict and proper use of citations so that the inclusion of other ideas in the dissertation are clearly distinguished,

3. that in the process of completing this doctoral thesis, no intermediaries were compensated to assist me neither with the admissions or preparation processes, and in this process,

- No remuneration or equivalent compensation were provided
- No services were engaged that may contradict the purpose of producing a doctoral thesis

4. that I have not submitted this dissertation or parts of this dissertation elsewhere.

I am aware that false claims (and the discovery of those false claims now, and in the future) with regards to the declaration for admission to the doctoral examination can lead to the invalidation or revoking of the doctoral degree.

____

Date, Signature

Appendix B

Declaration of Co-authorship

The following table outlines my contribution with respect to two of the three papers included in this dissertation (that is, chapters 2, and 4). Note that I am the sole contributor to chapter 3 ("**Stochastic debt sustainability analysis using time-varying fiscal reaction functions – An agnostic approach to fiscal forecasting**"), which is a single-author paper.

TABLE B.1: Declaration of Co-authorship

Chapter 2	**Fiscal prudence: It's all in the timing – Estimating time-varying fiscal policy reaction functions for core EU countries**, joint with Tino Berger and Ruben Schoonackers
Idea and conceptional work	contributed
Literature work	contributed
Data work	contributed
Empirical work	leading
Writing	contributed substantially
Chapter 4	**Government spending effects on the business cycle in times of crisis**, joint with Tino Berger
Idea and conceptional work	contributed substantially
Literature work	leading
Data work	leading
Empirical work	leading
Writing	leading

Tore Dubbert

Date