Université de Strasbourg

Ecole Doctorale Augustin Cournot ED 221

Bureau d'économie Théorique et Appliquée (BETA), UMR CNRS 7522

# Thèse

pour l'obtention du titre de Docteur en Sciences Economiques

*Présentée et soutenue publiquement le 18 janvier 2022 par*

**Kéa Baret**

# New empirical insights on fiscal rules performance

Co-directeur : **Amélie Barbier-Gauchard,** Maître de conférence HDR, Université de Strabourg

Co-directeur : **Théophilos Papadimitriou,** Professeur, Université de Thrace

## Jury

| | | |
|---|---|---|
| **António Afonso,** | Professeur, Université de Lisbonne | Rapporteur |
| **Etienne Farvaque,** | Professeur, Université de Lille | Rapporteur |
| **Belen Zinni,** | PhD, Head of Productivity Unit at OECD | Examinateur |
| **Xavier Debrun,** | PhD, First Advisor at National Bank of Belgium | Examinateur |
| **Alexandru Minea,** | Professeur, Université d'Auvergne | Examinateur |
| **Periklis Gogas,** | Professeur, Université Démocrite de Thrace | Examinateur |
| **Amélie Barbier-Gauchard,** | Maître de conférence HDR, Université de Strasbourg | Co-directeur |
| **Théophilos Papadimitriou,** | Professeur, Université Démocrite de Thrace | Co-directeur |

*L'université de Strasbourg n'entend donner aucune approbation, ni improbation aux opinions émises dans cette thèse; elles doivent être considérées comme propres à leur auteur.*

# Remerciements

Si, en se lançant dans une thèse, on s'imagine n'en jamais voir le bout, il est pourtant des personnes que l'on croise dans cette aventure qui nous permettent de mener à bien ce grand défi. Mais quelle tâche que d'écrire ces remerciements, partagée entre l'envie pressante de remercier toutes les personnes qui ont comptées et l'angoisse de n'oublier personne. Toutefois il y a des rencontres, des collègues, des amis qui marquent et que l'on ne peut oublier.

Ainsi, j'aimerais remercier tout d'abord mes directeurs de thèse qui m'ont accompagnée tout du long avec mes qualités et mes défauts: Amélie Barbier-Gauchard et Théophilos Papadimitriou. Ce combo parfait, alliant à la fois la politique budgétaire (Amélie) et une grande expertise du Machine Learning (Théophilos), m'a permis d'avancer et progresser dans ces domaines pourtant peu liés au départ, offrant ainsi une mixture originale grâce à leur profond investissement. Cette collaboration fut encore plus longue avec Amélie puisqu'elle m'accompagnait déjà dans mon mémoire de Master, co-dirrigé avec Alexandru Minea que je tiens aussi tout particulièrement à remercier ici. Non seulement il a motivé et soutenu l'idée de mon mémoire de recherche, mais il a aussi grandement contribué à son amélioration au point d'en faire le premier chapitre de cette thèse et une publication que nous partageons avec Amélie. Je lui dois également la découverte de la Roumanie ainsi que de nombreuses rencontres inoubliables, notamment le brillant Maxime Menuet et sa gentillesse.

Que l'on me permette ainsi de remercier ici tous les autres membres du jury, Etienne Farvaque et António Afonso qui m'ont fait l'honneur de rapporter cette thèse, ainsi que Belen Zinni, Xavier Debrun et Periklis Gogas qui ont accepté d'être de la partie.

Je remercie particulièrement Xavier Debrun pour sa bienveillance, son aide et d'avoir été le moteur de mon intérêt pour les règles budgétaires à la lecture de ses travaux. Il a été à la fois un exemple, un mentor et aura toujours mon admiration. Je remercie Antonio Afonso grâce à qui j'ai pu découvrir Lisbonne et rencontrer de nombreuses personnes formidables. Je remercie très profondément Belen Zinni avec qui j'ai partagé une année riche à l'OCDE en tant que stagiaire. Je la remercie de m'avoir choisie pour ce stage mais aussi pour toute la formation qu'elle m'a apportée ainsi que son humanité profonde et sa gentillesse débordante. J'en profite aussi pour remercier mon collège anglophone Ashley Ward dont les relectures ont été précieuses. J'adresse une mention particulière à ma division Trade and Productivity Statistics de l'OCDE avec qui j'ai partagé des moments professionnels et humains irremplaçables, et qui a rendu cette année difficile plus confortable.

Ce travail de recherche a été écrit en tant que doctorante au sein du Bureau d'Economie Théorique et Appliquée (BETA) que je tiens également à remercier pour m'avoir permis de réaliser cette aventure avec des expériences en colloques inespérées. Je remercie mes collègues qui m'ont soutenue durant ces trois années: Kenza Bousedra mon pillier du bureau 133, Pierre Pelletier, Stefano Bianchini pour avoir suscité mon intérêt pour le Machine Learning ainsi que Moise Sidiropoulos qui m'a si bien enseigné la macroéconomie et sans qui la rencontre avec Théophilos n'aurait pu être possible.

Je ne pourrais terminer ces remerciements sans mentionner ma famille. Je remercie mon conjoint Thibault qui m'a aimée, soutenue, accompagnée de Nice à Strasbourg, de Strasbourg à Paris... et qui m'a subie (aussi) pendant cette thèse. Comment poursuivre sinon en adressant mon affection à ma mère pour son soutien sans faille, peu importe les circonstances. Cette femme si forte, que j'ose espérer rendre fière, est une source d'inspiration au quotidien. Je remercie aussi ma petite soeur adorée que je soutiendrai à mon tour dans sa propre aventure de doctorante (je l'espère), mes grand-pères Richard et Pascal pour leur générosité de coeur, mes beaux-parents Brigitte et Patrick pour leur prévenence, ma belle-soeur claire pour ses lectures attentives. Je me dois également d'adresser quelques mots pour mes amies de toujours qui n'ont jamais cessé de croire en moi: Léna, Chloé, Diane et Marjorie. Il y a aussi mes rencontres strasbourgeoises et parisiennes, notamment Zohra pour nos discussions endiablées sur la dette et la politique monétaire, ainsi que Séverine pour nos inombrables heures de sport passées ensemble pour décompresser de ce dur travail. Aussi étrange que cela puisse paraître, j'ai aussi une pensée toute particulière pour mes deux chiennes qui m'ont longuement accompagnée dans ma rédaction, d'autant plus pendant le confinement. Leur présence fut le réconfort que nul humain n'aurait pu m'apporter.

Enfin, il me serait impossible d'oublier mon parrain, rejoint par ma grand-mère partie trop tôt pour me voir terminer cette thèse et dont nos discussions, la beauté et l'esprit raffiné me manquent aujourd'hui comme tous les autres jours.

**List of acronymes**

**BBR** = Budget Balance Rules

**ER** = Expenditure Rules

**RR** = Revenue Rules

**DR** = Debt Rules

**FR** = Fiscal Rules

**SGP** = Stability and Growth Pact

**p.p** = percentage point

**GFPI** = Global Fiscal Performance Index

**CAPB** = Cyclically Adjusted Primary Balance

**PSM** = Propensity-Score Matching

**SVM** = Support Vector Machine

**LASSO** = Least Absolute Shrinkage and Selection Operator

**RBF** = Radial Basis Function

**DML** = Double/Debiased Machine Learning

**OECD** = Organisation for Economic and Cooperation Develoment

**ONS** = Office for National Statistics

**IMF** = International Monetary Fund

**EU** = European Union

**EMU** = Economic and Monetary Union

**OCA** = Optimal Currency Area

**UK** = United Kingdom

**CFC** = Consumption of Fixed Capital

**GFCF** = Gross Fixed Capital Formation

# Summary

*Note: Le résumé de thèse est rédigé en Français selon le format exigé par l'Ecole doctorale Augustin Cournot à laquelle le candidat est rattaché. Le reste de la thèse est intégralement rédigé en Anglais.*

*The thesis summary is written in French according to the format required by the Augustin Cournot Doctoral School to which the candidate is affiliated. The rest of the thesis is written in English.*

Les discussions sensibles autour de la dette publique et du déficit public ont traversé les siècles, notamment en France où certains rois avaient pour coutume de financer leurs guerres par emprunt. Cette tendance à la dépense publique excessive a notamment mené la France jusqu'à la Révolution française de 1789 puisque les Français n'avaient, si l'on en croit Rousseau (1782, 1789), plus de pain pour se nourrir tant l'impôt était élevé et, pourtant, insuffisant pour couvrir les dépenses de la royauté. Deux siècles plus tard, les deux guerres mondiales (1914-1918, 1939-1945) ont creusé de manière spectaculaire les dettes des pays ayant pris part aux conflits. La dette publique est cependant associée à de nombreux autres mouvements sociaux dans le monde et ces considérations se sont accrues dans une période plus récente avec la crise des subprimes (2008-2009), la crise des dettes souveraines (2010-2012) et la crise pandémique (2020-2021) qui ont à nouveau accentué la détérioration des finances publiques. Une mauvaise gestion des finances publiques peut donc être un facteur d'instabilité politique et sociale et, c'est en particulier l'héritage de cette dette pour les générations futures qui fait débat. C'est ce que Wyplosz (2012) décrit comme le problème du « pool des communs » ou, autrement dit, dans quelle mesure est-il juste de charger nos générations futures pour nos souhaits présents ? De nombreuses analyses, notamment les modèles théoriques dits à générations imbriquées (intergénérationnels), comme dans le travail de Burbidge (1982), ont mis en évidence des effets négatifs liés à l'héritage des dettes publiques, pour le bien-être

des nouvelles générations. Outre l'instabilité politique et sociale, c'est aussi une instabilité macroéconomique qui pèse parmi les conséquences potentielles de la dette publique, comme le suggère notamment « l'effet boule de neige » mis en évidence par Pigou (1929). Ces risques de déséquilibre macroéconomique apparaissent possiblement encore plus dramatiques pour les unions monétaires comme la zone euro, si l'on se réfère à Théorie des Zones Monétaires Optimale de Mundell (1963). Les niveaux grandissants des dettes publiques peuvent donc être une menace pour la soutenabilité des finances publiques. Dans l'Union Européenne (UE), où l'une des conditions d'accès au statut d'état membre fixe le niveau maximal d'endettement à 60% du PIB selon les critères du Traité Maastricht (1992), la dette publique a avoisiné en moyenne 74,6% entre 2000 et 2019, certains pays dépassant parfois 100% comme la France ou l'Italie depuis la crise pandémique liée à la Covid-19 (2020). Ainsi, assurer la discipline budgétaire apparait donc comme une politique inévitable afin de garantir la stabilité de l'ensemble de l'économie. Le concept de discipline budgétaire concerne l'ensemble des moyens mis en œuvre afin de garantir des finances publiques saines et une dette publique soutenable. Parmi les outils pouvant être employés afin mettre en oeuvre la discipline budgétaire, les règles budgétaires numériques sont apparues dans les années 90 et ont connu une évolution grandissante aussi bien dans leur implémentation que dans leur « design ». Cette émergence des règles budgétaires ne fait pas exception dans le cadre de l'UE qui a expérimenté les premières règles supranationales avec la formalisation des critères de convergence du Traité de Maastricht (1992) dans le Pacte de Stabilité et de croissance (1997). Les règles budgétaires numériques se définissent comme des contraintes imposées sur les indicateurs des finances publiques et peuvent donc concerner des règles de budget/solde publique (lorsqu'elles concernent la balance/déficit publique), des règles de revenu (lorsqu'elles concernent les taxes), des règles de dépense ou encore des règles de dette. Toutefois, cette définition ne saurait renseigner sur l'efficacité de ces règles, c'est pourquoi Kopits et Symansky (1998) ont proposé une liste de propriétés que doivent revêtir les règles budgétaires pour être de « bonnes règles » et assurer leur rôle efficacement. Une règle idéale serait donc claire, correctement définie vis-à-vis de l'indicateur qu'elle vise, en adéquation avec les politiques économiques mises en œuvre, simple, contraignante, soutenue par des mesures et moyens adjacents, flexible. La tâche n'est donc pas aisée et, Debrun et Jonung (2019) ont mis en avant l'existence d'un triangle d'impossibilité propre aux règles budgétaires, tout comme il existe un en politique monétaire pour les régimes de change (voir Mundell (1982)). Selon les auteurs, il apparaît impossible pour une règle d'être à la fois simple, flexible et contraignante. Les règles sont ainsi régulièrement débattues et critiquées, peu importe leur niveau d'application (sous-national, national, supranational) parfois même au point d'envisager leur suppression comme envisagé par Blanchard et al. (2021). Dans ce contexte, il apparaissait donc nécessaire de tenter de répondre à une problématique importante, à savoir « Les règles budgétaires sont-elles efficaces ? ». Cette thèse a ainsi pour but d'analyser la performance des règles budgétaires à travers des apports empiriques afin d'apporter un éclairage sur l'efficacité des règles budgétaires, dans un contexte de débat houleux concernant leur avenir. La performance des règles budgétaires renvoie à différents concepts qui ne sont pas mutuellement exclusifs, à savoir: leur effet disciplinant sur le comportement des gouvernements, leur rigueur (souvent mesurée par des indicateurs de « force ») ou sur la base de leur indice de respect, mais également leur impact macroéconomique qui peut faire référence à des capacités stabilisatrices ou un effet sur des indicateurs sociaux.

Par conséquent, une partie de la littérature considère que l'efficacité des règles budgétaires est mesurée par leur effet disciplinant sur le comportement budgétaire des gouvernements. Dans cette littérature, la performance des règles budgétaires concerne la capacité des règles budgétaires à assurer la discipline budgétaire. Chez Bohn et Inman (1996), une attention particulière est accordée à l'impact des règles budgétaires sur les indicateurs budgétaires reposant sur le solde public ou la dette publique. Debrun et Kumar (2007) ont révélé l'absence de réponse significative du solde primaire corrigé des variations cycliques (cyclically adjusted primary balance en anglais ; CAPB par la suite) et leurs résultats sont cohérents avec les conclusions d'Escolano et al. (2012) sur le groupe des pays de l'UE-15. Au contraire, la réponse du CAPB est significative et positive dans les études de Debrun et al. (2008) et Marneffe et al. (2010). Dans la même lignée, Foremny (2014), Badinger et Reuter (2015) ou encore Bergman et al. (2016) ont conduit le même type d'analyse et trouvé des résultats significatifs.

La performance des règles est un large concept qui renvoie aussi à la crédibilité (critère clef de Kopits et Symansky (1998)) et au respect des règles, faisant notamment le lien avec la « surveillance des règles », afin de veiller à leur bonne mise oeuvre/application. Ainsi, la Commission Européenne et le Fond Monétaire International (FMI) disposent de bases de données sur les règles budgétaires et leurs caractéristiques. Ces dernières permettent d'évaluer leur force qui porte notamment sur leur inscription dans la loi nationale, leur indice de rigueur, leur niveau de flexibilité, la présence de sanctions pour leur non-respect... Larch et Santacroce (2020) ont ainsi proposé le « EU supranational fiscal rules compliance tracker » et ont identifié les déterminants de la conformité aux règles budgétaires supranationales du Pacte de Stabilité et de Croissance (PSC). Ils ont montré qu'entre 1998 et 2019, le respect des règles budgétaires inclues dans le PSC semble être lié à des variables macroéconomiques clés telles que les événements de politique budgétaire pro-cyclique, les institutions indépendantes de surveillance (faisant référence aux « chiens de garde budgétaires » (« independent fiscal watchdogs ») décrits par Debrun et Jonung (2019)) et la qualité de la gouvernance (tel que l'indice d'efficacité du gouvernement construit par la Banque Mondiale). Ainsi de nombreuses études se sont focalisées sur les déterminants du respect des règles. Parmi elles, Reuter (2019) a travaillé sur l'UE de 1995 à 2005, en utilisant une fonction logistique et a trouvé que les caractéristiques particulières des règles budgétaires telles que leur base légale, la présence d'un organe de contrôle indépendant, le degré de fragmentation du gouvernement, et le cycle politique sont corrélés avec la conformité aux règles budgétaires nationales. En effectuant une analyse similaire dans les économies d'Afrique sub- sahariennes, Nandelenga et Ellyne (2020) ont constaté que le PIB, les paiements d'intérêts de la dette ou encore les cycles électoraux ne semblent pas avoir d'impact sur le respect des règles budgétaires nationales. Les déterminants du respect des règles apparaissent donc comme hétérogènes, dépendant à la fois du contexte d'étude mais aussi du type de règle considéré.

La performance des règles peut aussi viser des indicateurs plus larges que la discipline budgétaire, ainsi ce concept touche à l'économie dans son ensemble avec des effets potentiellement stabilisateurs sur l'activité économique. Notamment Sacchi et Salotti (2015) et Guerguil et al. (2018) ont montré que les règles nationales avaient un pouvoir de stabilisation sur les variations du PIB. Larch et al (2021) ont fourni des résultats similaires concernant l'effet des règles supranationales. Ainsi la performance des règles concerne la croissance du PIB, mais également de nombreux autres indicateurs macroéconomiques parmi lesquels des indicateurs de bien-être social, comme par exemple les inégalités sociales, puisque Hartwig et Strum (2019) ont montré que les règles budgétaires étaient liées à davantage d'inégalités dans l'UE, et Combes et al. (2019) ont mis en évidence que les règles de dépenses pouvaient accroître les inégalités dans les pays en développement mais cet effet n'est pas vérifié pour les règles de budget et de dette. Ainsi, leurs effets peuvent couvrir un large panel d'indicateurs, et le signe de cette causalité reste indéfini, suggérant qu'il est possible que ces effets puissent aussi être de nature indésirable. Il serait en effet naïf pour un décideur politique de considérer les règles budgétaires sans s'intéresser aux contreparties des effets bénéfiques qu'elles peuvent engendrer. Il serait tout autant déraisonnable d'écarter complètement les règles en ne considérant que leurs effets pervers. Une investigation empirique apporte ainsi la possibilité de mesurer la performance des règles, ainsi que ses limites.

L'importance de « mesurer » fait référence au choix des modèles retenus mais aussi à la qualité des indicateurs ciblés par les règles qui peuvent conduire à des erreurs de jugement sur la performance des règles si ceux-ci sont initialement calculés de manière insuffisamment précise. L'ensemble des études susmentionnées ne sont pas intéressées à la qualité de mesure des indicateurs ciblées par les règles, alors qu'ils sont pourtant un facteur important susceptible d'affecter la mesure de la performance des règles.

Cette thèse offre donc une analyse empirique de la notion de performance des règles dans le but d'identifier les canaux par lesquels elles opèrent ainsi que leurs limites, afin de les définir/construire judicieusement. Bien que les études empiriques sur les règles puissent s'avérer nombreuses, leurs résultats sont à la fois controversés et parfois opposés. Ce constat résulte notamment d'une insuffisante prise en compte des biais statistiques du domaine de l'économétrie et qui accompagnent l'analyse des règles, comme mis en évidence par Heinemann et al. (2018). Cette thèse propose donc l'utilisation de modèles économétriques robustes afin de fournir des conclusions interprétables et généralisables. Elle offre également de nouvelles perspectives empiriques grâce à l'utilisation du Machine Learning, dont les performances sont mises en évidence de façon grandissante aussi bien dans le domaine de la prévision que de l'analyse causale, mais dont les applications sont rares dans le domaine des politiques publiques, et davantage encore sur des problématiques spécifiques à la discipline budgétaire. Enfin, nous n'avons pas connaissance, à ce jour, de l'existence d'une analyse empirique mettant explicitement en exergue l'importance des méthodologies statistiques utilisées dans les comptes nationaux des pays afin de calculer les indicateurs cibles des règles, et dont la précision apparaît indispensable pour la bonne évaluation de la performance des règles.

**Le premier chapitre** de la thèse s'inscrit donc dans le premier volet de la littérature précédement évoquée et s'attèle donc à évaluer empiriquement la performance des règles budgétaires et notamment leur capacité à assurer la discipline budgétaire. En effet, après chaque grande crise économique de ces dernières décennies, les règles budgétaires ont été questionnées. Dans ce contexte de vif débat, ce chapitre propose de revenir au rôle premier des règles budgétaires qui est d'assurer la discipline budgétaire. Le champ des sciences politiques des années soixante-dix, et notamment Buchanan, a rappelé que « Les gens, y compris les politiciens, sont principalement motivés par leur propre intérêt » (voir la théorie des « Choix Publics » avec Buchanan et Tullock (1962) ou encore Buchanan et Wagner (1977)). Cela suggère donc que le problème du pool des commun (Wyplosz (2012)) ou encore le problème d'incohérence temporelle (Kydland et Prescott (1977)) sont importants et font planer une menace sur la soutenabilité des finances publiques. Les règles budgétaires s'imposent donc comme un instrument afin de lier les mains des décideurs publiques tout comme les règles d'inflation lient les mains des banquiers centraux. Si tentait qu'elles assurent encore ce rôle, le maintien des règles budgétaires pourrait ainsi être défendu. Ainsi ce chapitre adresse la question suivante : « Les règles budgétaires permettent-elles d'assurer la disciplinaire budgétaires ? ». Cette analyse empirique requière donc le choix d'un estimateur robuste afin de garantir des résultats non discutables. Cette analyse étant conduite dans le contexte spécifique de l'Union Européenne, la prise en compte des hétérogénéités entre pays membres apparaît comme un challenge de taille. De plus, et par définition, la discipline budgétaire elle-même n'est pas directement mesurable et est approximée par des variables dites de proximité. Ces dernières sont susceptibles d'engendrer des biais d'endogénéité bien connus en économétrie, comme le biais de causalité inverse ou le biais d'omission. Le premier chapitre propose donc une étude économétrique au travers d'un modèle de propensity-score Matching qui lutte contre les biais susmentionnés susceptibles d'accompagner l'étude de la performance des règles budgétaires. L'approche du premier chapitre évalue ainsi l'impact des règles budgétaires nationales sur la discipline budgétaire (alternativement mesurée par le solde public structurel primaire (CAPB) et par un nouvel indicateur composite de performance budgétaire globale né de ce chapitre) dans l'UE entre 2000 et 2013. Robuste à de nombreux tests, spécifications et estimateurs alternatifs, ce chapitre a notamment permis de montrer que les règles ont un effet significatif sur la discipline budgétaire mais les conclusions dépendent du type de règle considéré, de facteurs structurels propres aux pays, et la mesure de discipline budgétaire retenue constitue également un facteur pouvant affecter les résultats.

Cette thèse ne néglige pas les lacunes de performance des règles notamment leur défaut d'application puisque le **chapitre 2** constate qu'entre 2006 et 2018, la règle supranationale de l'UE -à savoir le Pacte de Stabilité et de Croissance (PSC)- n'a été respectée que dans 50% des cas, et ce constat est également valable pour les règles nationales en général (voir Reuter (2019) ou Caselli et al. (2018)). Le chapitre 2 s'intéresse donc à la problématique suivante : « comment prévoir le respect des règles budgétaires afin de renforcer leur crédibilité et leur performance ? ». Pour y répondre, le deuxième chapitre étudie les 28 pays de l'UE entre 2006 et 2018, et se concentre sur la prédiction des cas de conformité au PSC avec une attention particulière portée à la règle des 3% visant le déficit public. Ce chapitre propose un modèle automatisé, employant des méthodes prévisionnelles de Machine Learning particulièrement précises concernant l'anticipation des situations de non-respect du PSC. Une stratégie à deux étapes est privilégiée avec dans un premier temps l'identification des déterminants du respect du PSC puis la prédiction du respect du PSC en utilisant les déterminant-clefs retenus à la première étape. La sélection des déterminants utilise un logistic-LASSO qui permet d'éliminer les variables non importantes pour la conformité au PSC et ne retenir que les meilleurs prédicteurs. A la seconde étape, la confrontation de plusieurs modèles de Machine Learning (à savoir différents modèles à vecteurs de support (communément appelé Support Vector Machine (SVM) en anglais) dont un à noyau linéaire, un noyau quadratique et un noyau RBF), comparés à une fonction logistique issue de l'économétrie standard, met en évidence que l'utilisation du modèle à vecteur de support à noyau linéaire, combiné au logistic-LASSO, apparaît être l'outil le plus performant afin de prévoir les cas d'échec dans la conformité des pays de l'UE à la règle supranationale. Cette approche en deux étapes permet aussi de quantifier les efforts nécessaires à mettre en oeuvre par les pays afin de garantir la conformité aux conditions prévues par le pacte. Les déterminants du respect du PSC peuvent servir d'outils d'ajustement afin de ramener les finances publiques à une situation dite de « conformité à la règle des 3% du PSC ». En effet, en appréhendant la distance de chaque pays par rapport aux conditions prévues par le PSC, le modèle permet d'envisager des solutions quantitatives afin de garantir l'application du PSC dans l'UE. Ce chapitre propose donc un outil de management du risque budgétaire en veillant à prévoir la conformité au

PSC afin d'assurer sa performance. Il met également en avant les avantages de l'utilisation du Machine Learning à des fins de prévisions puisqu'il surpasse l'économétrie standard dans ce cas d'étude.

Ensuite, le **chapitre 3** analyse les implications sous-jacentes à l'application des règles budgétaires, puisque des effets négatifs non voulus pourraient limiter leur performance. Le chapitre 3 tente de répondre à la question suivante : « dans quelle mesure le respect règles budgétaires peut-il impliquer des effets dommageables pour le bien-être social ? ». Ce troisième chapitre utilise un modèle récent d'inférence causale issu du Machine Learning (Double/Debiased Machine Learning développé par Chernozhuckov et al. (2018)) afin d'évaluer les effets pervers d'une stricte application des règles sur le bien-être social. Ce chapitre porte une attention particulière à la définition de « conformité aux règles » en considérant une définition « stricte » reflétant uniquement si le pays a dépassé ou non la limite fixée par la règle ; mais aussi une définition à la fois plus sophistiquée et plus flexible qui tient compte de la présence de clauses échappatoires (autorisant les pays à dévier de la règle en situation de conjoncture économique exceptionnellement critique). En se focalisant sur le cas particulier des règles nationales de budget, ce modèle en deux étapes identifie dans un premier temps les déterminants du respect des règles nationales portant sur le solde budgétaire puis évalue l'effet du respect de ces règles sur de nombreux indicateurs faisant le lien avec le bien- être social (parmi lesquels notamment l'investissement public productif, les dépenses sociales et des indicateurs d'inégalités sociales). La première étape consiste en effet à isoler les variables affectant à la fois le respect des règles et chacun des indicateurs de bien-être social, afin d'extraire ces informations et éviter un biais d'endogénéité. Cette procédure de sélection est réalisée avec un LASSO (déjà employé dans le deuxième chapitre) ainsi qu'un modèle de « l2- boosting » (à titre de test de robustesse) qui permettent de tester un large nombre de candidats potentiels parmi les déterminants potentiels du respect des règles et les variables explicatives du bien-être social. Les deux estimateurs fournissent des résultats convergents en sélectionnant les mêmes prédicteurs et montrent notamment que le respect des règles dépend de nombreux facteurs comme les préférences des électeurs, confirmant ainsi – comme le suggérait Wyplosz (2012) - l'importance de tenir compte des problèmes d'endogénéité lorsque l'on évalue empiriquement la performance des règles budgétaires. La deuxième étape consiste à évaluer l'effet du respect des règles sur les indicateurs de bien-être social, et est obtenue par la combinaison d'une régression consistant à orthogonaliser les résidus de la première étape (similaire à une approche à la Frisch-Waugh-Lovell (Frisch et Waugh (1933), Lovell (1963)) avec une validation croisée. Ces procédures permettent d'exclure le biais d'endogénéité puisque les informations contenues par les prédicteurs affectant à la fois le traitement (le respect des règles) et la variable dépendante (les indicateurs sociaux retenus dans l'étude) sont extraites permettant ainsi d'isoler l'effet du respect des règles. Ce troisième chapitre montre finalement que la « stricte application » des règles budgétaires (définition large de la conformité) ne se fait pas à n'importe quel prix puisqu'un strict respect des règles budgétaires peut avoir des effets négatifs sur les dépenses sociales. En effet l'utilisation d'une définition du respect des règles reflétant uniquement si le pays a ou non dépassé la limite imposée (et ne tenant donc pas compte de la possible existence de clauses échappatoires) affecte négativement la dépense de consommation finale du gouvernement qui comporte des dépenses sociales comme les transferts sociaux. Nous observons également que ce respect, au sens strict, des règles budgétaires accroît les inégalités sociales mesurées par l'indice de Gini, avec un impact plus important sur les classes les plus pauvres. La prise en compte des clauses échappatoires permet de nuancer les effets négatifs qui vont avec les règles puisque cette flexibilité dans la notion de « respect de la règle » introduit un effet positif sur la croissance économique. Ainsi, un respect des règles en adéquations avec le cycle économique permet de préserver les objectifs économiques mais ne permet toutefois pas de limiter les effets sur les indicateurs de bien-être social. Il semblerait en effet que le respect des règles, peu importe la définition retenue du respect de la règle, touche à des choix stratégiques des gouvernements dans la composition des dépenses publiques, renvoyant plutôt à la part structurelle des politiques économiques.

Un autre élément important qui pourrait impacter la performance des règles budgétaires concerne la qualité de la mesure des indicateurs cibles. Ce point n'a pas encore été soulevé par la litérature existance, pourntant, si les indicateurs visés par les règles numériques n'étaient pas mesurés de la façon la plus optimale possible, l'évaluation de la performance des règles pourrait être affectée. Le **chapitre 4** pose donc la question suivante : « dans quelle mesure les méthodes statistiques utilisées pour calculer les indicateurs cibles des règles peuvent impacter leur performance ? ». Cette question est d'autant plus importante que le débat sur la pertinence des règles s'accentue face au besoin croissant d'investissements publiques dans de nombreux pays développés et en développement, aussi bien pour se remettre des crises économiques des dernières décennies que pour assurer un rattrapage économique pour les plus en retard. Ce besoin est pourtant accompagné de dettes publiques parfois colossales, et l'arbitrage entre « besoins en investissement » et « discipline budgétaire » place les règles dans une position délicate. Ce débat remonte notamment aux années 2000 où déjà la littérature sur la règle d'or des finances publiques posait les prémisses de l'importance de la flexibilité des règles, notamment vis-à-vis de l'investissement public (voir notamment Creel et Saraceno (2010), Huart (2012), Creel et al. (2014)). Cette règle d'or, dont l'objectif cible est le solde hors investissement public net afin de libérer la contrainte budgétaire de long terme pesant sur l'investissement, est aujourd'hui une solution sérieusement envisagée (d'après le discours du commissaire européen Paolo Gentiloni (2021)). Les discussions sur la future réforme des règles budgétaires dans le cadre de l'UE semblent refléter une volonté d'introduire des marges de manœuvres sur les dépenses publiques productives dans l'UE. Le dernier chapitre de cette thèse propose ainsi une analyse de la sensibilité de la cible de la règle d'or -à savoir le solde public hors investissement public net- aux méthodes statistiques employées pour mesurer l'investissement public net. Cette analyse utilise le Royaume-Uni comme cas d'étude entre 1998 et 2016. Ce choix est notamment motivé par le fait que le Royaume-Uni a longtemps eu une règle d'or et les données utiles au calcul de l'investissement public net sont disponibles. Le dernier chapitre met finalement en évidence que les changements dans la méthode de calcul de l'investissement public net, peuvent affecter significativement la cible de la règle d'or et donc, indirectement, sa performance. Cette cible dépend des méthodes statistiques qui permettent de dériver la consommation fixe de capital (aussi appelée dépréciation) qui est déduite de l'investissement brut afin d'obtenir l'investissement net. En effet, la consommation fixe de capital est obtenue à partir des mesures du stock net de capital qui peut être calculé en introduisant les flux passés d'investissement (formation brute de capital fixe) dans ce que l'on appelle « la méthode de l'inventaire perpétuel » (voir OCDE (2009)). Il est nécessaire pour cela de disposer de séries longues et détaillées sur la formation brute de capital fixe (publique dans notre cas) et de définir une fonction d'âge-prix (qui permet d'obtenir le stock net de capital) ainsi que ses paramètres comme la durée de vie des actifs et le taux de dépréciation des actifs. Parmi les fonctions d'âge-prix nous testons l'approche géométrique puisqu'il s'agit d'une approximation raisonnable pour une cohorte d'actifs, comme suggéré par l'OCDE (2009), et nous comparons les résultats à ceux publiés par l'OCDE qui reportent les données officielles de l'Office for National Statistiques (ONS) du Royaume-Uni. Également, nous tenons compte du fait que l'ONS a introduit de nombreuses modifications méthodologiques en 2019, concernant notamment la durée de vie des actifs (affectant donc leur vitesse de dépréciation/taux de dépréciation) ainsi que l'approche pour obtenir le stock net de capital. Avant 2018, l'approche de l'ONS consistait à combiner une fonction d'âge-prix « straight- line » avec une fonction « normale » de retirement des actifs (afin d'éliminer les actifs en fin de vie disparus durant le processus) afin d'obtenir le stock net de capital et dériver la dépréciation du capital public. Nous reproduisons donc les anciennes séries de l'ONS et utilisons à titre de comparaison les anciennes séries officielles. Toutes ces comparaisons permettent d'évaluer l'impact des changements dans la forme des fonctions d'âge-prix sur la cible de la règle d'or. En fixant ensuite la fonction d'âge-prix (en utilisant une forme géométrique), nous faisons varier le taux de dépréciation des actifs pour analyser l'impact d'un changement des paramètres des fonctions d'âge-prix sur la mesure de la performance de la règle. Les changements dans les hypothèses faites sur les fonctions qui permettent d'obtenir la consommation de capital fixe publique, ainsi que sur leurs paramètres, sont cruciales car elles peuvent modifier à la fois l'investissement public net mais aussi le PIB puisque son calcul prend en compte la consommation de capital fixe (CCF). Les principaux enseignements montrent que les changements dans la fonction d'âge-prix engendrent des variations dans les séries de la consommation fixe de capital atteignant jusqu'à 0.75% du PIB et qui se reportent donc sur l'investissement public net puis sur la balance publique excluant l'investissement public net. La modification des taux de dépréciation génère aussi des impacts sur la CCF mais plus faibles, n'excédant pas 0.2% du PIB. Le niveau de détails des calculs semble aussi avoir une importance majeure : calculer la CCF à partir de données très détaillées -à savoir au niveau des industries et actifs publics, par rapport au niveau plus agrégé des actifs publics sans tenir compte des industries/sous-activités publiques- implique des variations de l'investissement public net pouvant aller jusqu'à 0.8% du PIB. Des erreurs méthodologiques pourraient ainsi conduire à des évaluations erronées de la règle d'or et fournir des recommandations de politiques économiques inadéquates. En effet, comment mesurer quantitativement les efforts budgétaires à fournir afin d'ajuster le solde public excluant l'investissement public net si celui-ci est mal calculé ? La mise en œuvre d'une potentielle réforme des règles budgétaires, notamment du PSC dans le cadre de l'UE, au profit d'une règle d'or exigerait alors un système de comptabilité national transparent et rigoureux dans chaque pays membre, ainsi qu'une harmonisation des pratiques statistiques dans le calcul des comptes nationaux des pays. Ces recommandations sont d'autant plus importantes que les impacts potentiels sur le calcul du PIB suggèrent des conséquences plus larges couvrant de nombreux indicateurs macroéconomiques comme la croissance potentielle, les calculs du cycle économique ainsi que les valeurs cibles de règles budgétaires exprimées en pourcentage du PIB.

Cette thèse s'intéresse donc à toutes les facettes du concept de performance des règles budgétaires. Elle apporte ainsi une vision à la fois nouvelle et élargie de la mesure de la performance des règles budgétaires afin d'éclairer les débats futurs autour d'une réorientation éventuelle des règles budgétaires existantes. Les enseignements ont d'abord suggéré que les règles ont encore leur place dans la politique budgétaire puisque leur performance vis-à-vis de la discipline budgétaire semble bien valide. Aussi, il est possible d'offrir des cadres simplifiés permettant de surveiller la bonne implémentation des règles et veiller à leur performance, notamment par l'utilisation de modèles prédictifs récents issus du Machine Learning. Toutefois, les résultats fournis permettent aussi d'identifier des effets pervers qui doivent servir d'enseignement afin d'offrir des solutions (dès la base) pour la construction des règles qui, bien que ne pouvant satisfaire tous les critères de Kopits et Symansky (1998), pourraient apparaître comme des « optima de second rang ». Ces dernières conclusions sont aussi issues d'un modèle de Machine Learning mais dans une approche de causalité, montrant encore une fois que ces techniques ont une place dans le domaine de l'analyse des politiques publiques. Aussi, il est important que les indicateurs ciblés par les règles soient eux-mêmes bien construits et donc bien calculés. D'importants efforts doivent être mis en œuvre afin de calculer certains indicateurs de comptes nationaux ciblés par les règles budgétaires et notamment la consommation fixe de capital qui permet de calculer l'investissement net qui est la cible de la règle d'or des finances publiques, très présente dans le débat sur la réforme des règles dans le cadre de l'UE. Enfin cette thèse offre de nombreuses pistes pour des recherches futures, comme de vérifier l'effet disciplinant de règles bien précises comme des règles flexibles (aussi appelées règles de seconde générations) ou dans un autre cadre que celui de l'UE (comme dans les pays de l'Union monétaire de l'Afrique centrale (UMAC) ou de l'Union monétaire ouest-africaine (UMOA)). Il est également possible de construire des modèles visant à prédire la conformité aux règles nationales ainsi qu'aux différentes règles incluent dans le PSC puisque l'exercice réalisé dans cette thèse s'est concentré sur la règle supranationale des 3% visant le déficit public inclue dans le PSC. Il serait aussi d'intéressant de vérifier les effets pervers sur le bien-être social d'autres règles budgétaires mais cela demande la construction de nouvelles bases de données sur les règles et/ou l'amélioration/extension de celles déjà existantes (la base de données du FMI s'arrêtant en 2015 et celle de la Commission Européenne se focalisant uniquement les pays membres de l'UE). Enfin, cette thèse ouvre la voie à des questions sur l'avenir des règles budgétaires : serons-nous en mesure de tirer leçons des enseignements empiriques sur la performance des règles ? Les institutions indépendantes en charge de leur surveillance et de leur mise en œuvre suffiront-elles à leur redonner de la crédibilité et assurer leur performance ?

# Contents

# General introduction

*This introduction is inspired by a co-authored chapter "Barbier-Gauchard, Baret and Debrun (2022), Government efficiency and fiscal rules, in Afonso, Jalles and Venâncio (Eds), Handbook on public sector efficiency, Edgar, forthcoming."*

"Toutes ces discussions morales et politiques ne seraient que des idées volantes si elles ne touchaient terre, et durement, sous la forme d'une question très concrète : qui va payer la dette publique française ? Il n'existe pas de petit-déjeuner gratuit, et tous ceux qui investissent ou consomment à crédit aujourd'hui utilisent l'épargne et le travail d'autrui. Si la croissance économique est insuffisante, l'impôt pèsera sur les générations futures qui paieront ce que nous consommons aujourd'hui." Perrot [2010]

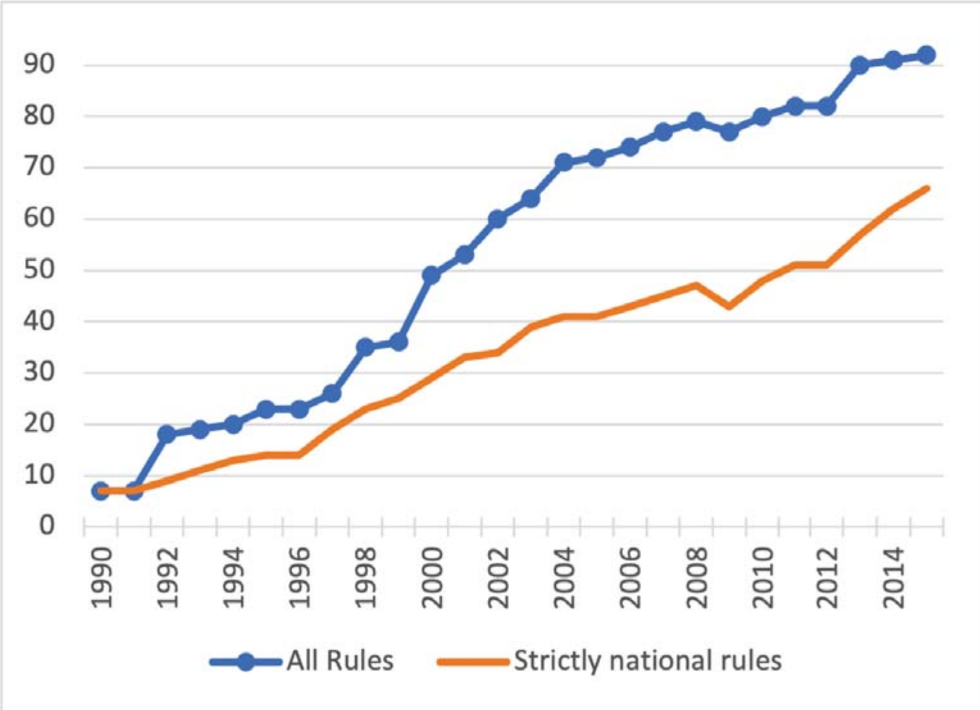
The issue of public debt has spanned centuries. By way of illustration, the French royal state has survived on credit since the beginning of the 16th century, especially to finance its wars. By the 17th century the situation was already so bad that Louis XIV, king of France, asked his banker, Samuel Bernard, to obtain credit, but ended up going bankrupt. In the 18th century, the public debt soared and caused widespread public discontent, leading to the "French Revolution of 1789". Two centuries later, new wars have widened the French deficit, and the problem had spread to other economies. Indeed, the two world wars (1914-1918 and 1939-1945) increased the risk of public debt unsustainability in all countries engaged in those wars, both allied and axis nations. Even though France is often considered a bad student in terms of public finance management, public debt exists the world over and the historic trend of growing public debt demands careful attention as countries continue indebting themselves. For example, the recent observation of the public balance in European Union (EU) member countries reveals an (unweighted) average of 2.4% deficit (expressed as a percentage of GDP) between 2000 and 2019, and was not once positive over the period. This tendency to run public deficit in EU member countries is accompanied by an important evolution of public debt which represents 74.6% of GDP on average between 2000 and 2019 compared with 60% at the birth of the EU in 1992. Focusing our interest on the evolution of public finance seems important and relevant since their mismanagement can have deep economic and social consequences. This pattern of increases in public debt has the potential to induce social revolts beyond those in France, such as in Poland with the *Solidarnosc*, or in Greece after the Global Financial Crisis (hereafter GFC) that imposed six fiscal recovery plans with negative consequences for social conditions. It appears that running excessive public deficit that feeds the public debt level is a major factor for political and macroeconomic (in)stability. The Snowball effect described by Pigou [1929] reinforced this finding and the inheritance of debts for future generations added the question of the "common pool" problem (Wyplosz [2012]). Phrased as a question: To what extent is it reasonable to charge present generations for the deficits of their elders? Overlapping models (Burbidge [1983] or Persson [1985]) have highlighted the potentially harmful effects of public deficit for the welfare of future generations. The economic crises of recent decades namely the GFC, the sovereign debt crisis and more recently the pandemic crisis due to COVID-19 once again raised the importance of regulating public finance. Such considerations matter even more in the context of monetary unions as discussed by Mundell [1963] who developed the Optimal Currency

Area theory.

Despite these discussions on the potential negative effects of public deficit and public debt, many countries still had high deficits, even before the pandemic crisis. We see at least three main arguments to explain government's behaviour regarding public deficit. First, when resources are not sufficient to cover all expenses, governments run a deficit to avoid defaulting on payments or imposing policies of austerity. Second, the neutrality of public debt addressed by Barro [1974] was challenged by at least two main findings: the wealth effect of feeding debt (Modigliani [1961]) and non-linear effects of public debt on economic growth, which suggest that there exists an optimal level of debt that maximizes economic growth (see Reinhart and Rogoff [2010] or Egert [2017]). Third, given the desire of governments to be re-elected, we observe a lack of government commitment to balancing public financing and paying down public debt, namely "temporal inconsistency" (see Kydland and Prescott [1977]). Consequently, monitoring public finance appears essential to ensure macroeconomic stability and fiscal discipline has received increasing attention as public deficits widened in many OECD countries.

Fiscal discipline relies on the mechanisms implemented by governments to achieve sound and sustainable public finance. Fiscal discipline should promote public finance stabilization when fiscal policy appears defective. In its annual activity report, the European Commission (Commission [2010]) defined the elements of the fiscal framework, which aims to ensure fiscal discipline. These elements include: annual numerical fiscal rules, independent institutions in charge of the monitoring and the proper implementation of fiscal rules, medium-term budgetary frameworks which are established over longer time horizons than numerical fiscal rules, and fiscal procedures. Fiscal discipline thus concerns these tools implemented to ensure a sustainable level of public debt, which is linked with the public balance. This thesis focuses on fiscal rules since their number has multiplied in the last three decades, a period during which they have become a point of contention and debate. It is possible to define a fiscal rule as a constraint set on public finance indicators in the form of a numerical target (Schwengler [2012]) that concerns public budget balance, public expenditure, public debt or public revenue. These targets may concern structural components, be expressed in nominal terms, or exclude certain types of investment... There is therefore a wide variety of possible targets for fiscal rules, with budget balance rules, in particular, offering numerous possibilities. The most prominent variants are: the cyclically adjusted primary balance (which corresponds to the public balance that exclude both interest payments on public debt and cyclical components) and the Golden rule (which belong to investment-friendly rules) that targets public balance excluding public net investment. Given their importance, the International Monerary Fund (IMF) and the European Commission regularly compile a worldwide dataset on fiscal rules in force. The dataset covers all types of numerical fiscal rules (budget balance rules, debt rules, expenditure rules and revenue rules) at all levels of government (central, regional and local, general government and social security). By constraining fiscal indicators, fiscal rules may help to achieve fiscal discipline, but they are not applied with the same degree of rigor, nor with the same degree of autonomy of governments subject to the rules. A key characteristic of fiscal rules concerns their "application level" which refers to the decision-making level that imposes fiscal rules (for instance, rules could be imposed by the national or supranational level). On the other hand, the "coverage level" refers to the decision-making level to which fiscal rules apply (i.e. rules could be imposed on subnational, national or supranational authorities). For example, there may be supranational rules that are imposed to a national level (this is the case of the Stability and Growth Pact -thereafter SGP- (1996) in the Euro Area). In that case, the "application level" is supranational while the "coverage level" is national. Conversely, in the case of a self-imposed fiscal rule, the "application level" is the same as the "coverage level": it is the decision-making level concerned which imposes rules on itself. Otherwise, when the degree of budgetary autonomy is limited by fiscal rules application at local level (as in the United States for states level), federal mechanisms are activated to assume budgetary functions. However, this mechanism is absent in the eurozone where there is no centralised mechanism for cushioning cyclical shocks. It should be noted, that fiscal rules may differ in many ways, especially in different geographic zones.

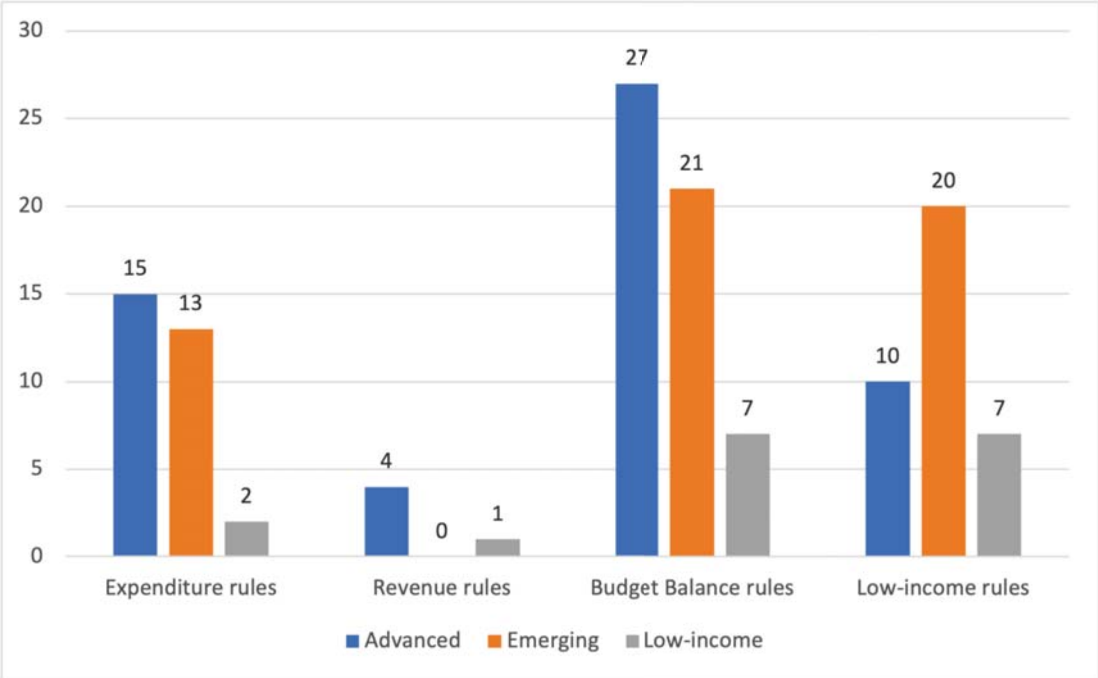
As shown in figure 1, in 2009, around 80 countries had national and/or supranational fiscal rules, compared with only 7 countries in 1990. According to Schaechter et al. [2012], in the same year (2009), more than 50 countries had national fiscal rules (including 20 in combination with supranational rules). This rapid expansion reflects the adoption of fiscal rules, in particular for European and Latin American countries, as well as the introduction of supranational rules, especially in low-income countries. Indeed, in 1992 the Maastricht Treaty introduced the first version of a supranational rule with the convergence criteria for EU accession and set in motion the implementation of supranational rules across the world such as the SGP (1996) for the Euro Area.




Source: authors using IMF Fiscal rules' Database.

Figure 1: **Number of countries with fiscal rules, in the world, between 1990 and 2015**

In 2009, the most popular fiscal rules relate to constraining public balance and public debt as illustrated in the Figure 2. Almost 80% of the developed countries used a fiscal rule that relates to the public balance or public debt. This may reflect a preference of governments for indicators closely linked to the public finance sustainability.




Source: authors using IMF Fiscal rules' Database.

Figure 2: **Number of national fiscal rules by type of rules and by countries' level of development, in the world, in 2015**

Following this overview, a burning issue remains unsolved: are fiscal rules performant? The crucial question of fiscal rules performance is the subject of heated debate. What is an efficient fiscal rule? What criteria should we focus on to assess fiscal rules performance? Does there exist a rule that can help to stabilize public finance while avoiding government spending bias? Is there any rule that has no side-effect, or is a performant rule only a "strong" rule or a "credible rule" in the sense that it is imposed in a "binding" way? Since the seminal work of Kopits and Symansky [1998], many subsequent works have attempted to assess the criteria for an "Ideal Fiscal Rule", but there is no consensus on this issue. Different approaches exist and fiscal rules performance may refer to several non-mutually exclusive concepts: i) with regard to their disciplining effect on government behavior, ii) with respect to their impact on macroeconomic variables, iii) regarding their rigor or strength proxied by fiscal rules strength indices, iv) on the basis of the degree of compliance (index) with the fiscal rules, or v) depending on their impact on Social field.

Even if Kopits and Symansky [1998] did not explicitly use the term "fiscal rules performance", their contribution constitutes the starting point of the fiscal rules' efficiency debate. Indeed, Kopits and Symansky (1998) identified eight properties with which an "ideal" fiscal rule must comply.

**Kopits and Symansky (1998)'s criteria**

1. Well definition of the rule regarding its target: to achieve effective enforcement;
2. Transparency: the rule should avoid opaque fiscal policy intention;
3. The rule should be "adequate with respect to the specified proximate goal";
4. General consistency: the rule must be "consistent" with the objectives of economic policy;
5. "Simplicity": the calculation of the target must be able to be done without requiring sophisticated calculation techniques;
6. Flexibility: governments must be able to continue to carry out their missions;
7. Enforceability: the rule should be credible including control procedures and the application of sanctions must be applied in an impartial and consistent manner;
8.The rule should be "supported by efficient policy actions": the rule should be included in a complementary fiscal framework.

The evaluation of different rules with regard to all of these properties leads to subjective judgments. How, for example, can we assess the degree of credibility of various rules which have so far never led to the application of sanctions? How to assess overall consistency? Nevertheless, a number of studies have tried to assess fiscal rules performance using the criteria of Kopits and Symansky [1998]. Among them, Buti et al. [2003] used a grid analysis and showed that the initial supranational fiscal rule of the SGP in the eurozone is better than a deficit rule excluding public investment, while Creel [2002] concluded the opposite. In this regard, Creel [2002] emphasises the subjective character of the values associated with each property. These analyses lead to an extended literature that proposed empirical assessment of fiscal rules performance.

Consequently, one strand of the literature considers that fiscal rules effectiveness is measured by their disciplinary effect on fiscal behavior. In this literature, fiscal rules performance concerns the ability of fiscal rules to achieve fiscal discipline. Following the paper of Bohn and Inman [1996], a particular attention is paid to the impact of fiscal rules on fiscal indicators relying on public balance or public debt. For example, Debrun and Kumar [2007] revealed the lack of a significant response of the Cyclically Adjusted Primary Balance (CAPB thereafter) when instrument variables are of to proxy fiscal rules, and their results are consistent with the conclusions of Escolano et al. [2012] on the group of EU15 countries. On the contrary, the response of CAPB is significant and positive in Debrun et al. [2008] and Marneffe et al. [2010] (see also Foremny [2014], Badinger and Reuter [2015] or Bergman et al. [2016]).

Nevertheless, there is no rule which can achieve all of the properties outlined by Kopits and Symansky [1998]'s properties. Similarly to the monetary policy trilemma developed by Mundell and Fleming in the 1960s, fiscal rules are also under a trilemma. According to Debrun and Jonung [2019], who highlighted the "fiscal rules trilemma", a fiscal rule cannot be flexible, simple and enforceable at the same time, and it is necessary to operate a trade-off between these different qualities. For example, if one rule favors simplicity, it is likely to lack enforceability. Several studies have pointed to a poor compliance track record, including Reuter [2019] who showed that EU governments are compliant with only 51% of national fiscal rules (similar findings are supported by Caselli and Reynaud [2019] for a larger panel of countries). In this landscape and facing this trade-off, how could countries achieve compliance? How can we increase the monitoring of fiscal rules? Can we identify compliance drivers and prevent failures in compliance? Questions such as these introduced a new literature on fiscal performance that focuses on fiscal rules compliance since the credibility of fiscal rules is one of the major characteristic of Kopits and Symansky [1998]. While the aforementioned studies have attempted to assess the compliance rate using the fiscal rules databases published by the European Commission or the IMF, as well as the EU supranational fiscal rules compliance tracker of Larch and Santacroce [2020], other studies identified the determinants of fiscal rule compliance. Among them, Reuter [2019] worked on the EU area from 1995 to 2005, using a logistic function and found that particular characteristics of the fiscal rules such as their legal basis, the presence of an independent monitoring body, the degree of government fragmentation, and the political cycle are correlated with compliance to fiscal rules. Conducting a similar analysis in sub-saharian economies, Nandelenga and Ellyne [2020] found that the GDP, the electoral cycles, levels of public debt and interest payments do not impact national fiscal rules compliance. Larch and Santacroce [2020] showed that between 1998 and 2019, compliance with fiscal rules included in SGP seem to be linked with key macroeconomic variables such as pro-cyclical fiscal policy events, the performance institutions (referring to the fiscal "watchdogs" described by Debrun and Jonung [2019]) and the quality of governance (such as the government effectiveness index constructed by the World Bank).

On the other hand, the assessment of fiscal rules also concerns their macroeconomic impact. Indeed, as previously discussed, public deficit may destabilize macroeconomic conditions and fiscal discipline may thus help to stabilize economies. There is thus a strand of literature considering that fiscal rules effectiveness should be measured by their impact on other macroeconomic variables. Indeed, as suggested by Bohn and Inman [1996], the fiscal rules should also be implemented for their potential stabilizing effects on macroeconomic aggregates. In that sense, attention is focused on reducing the volatility of macroeconomic indicators such as activity, employment, inflation or public expenditure. For example, Sacchi and Salotti [2015] underlined that, when strict fiscal rules are introduced, discretionary policy becomes output stabilizing rather than destabilizing (in particular by using balanced budget rules, rather than expenditure, revenues, or debt fiscal rules). However, they found that fiscal rules may not be able to affect inflation since inflation is under the responsibility of central bankers who are not necessarily confounded with fiscal policy makers. Guerguil et al. [2017] assessed the impact of different types of flexible fiscal rules on the procyclicality of fiscal policy, and found that investment-friendly rules reduce the procyclicality of both overall and investment spending. The effect appears stronger during bad economic periods and when the rule is enacted at the national level. Their results also showed that the introduction of escape clauses in fiscal rules does not seem to affect the cyclical stance of public spending. The inclusion of cyclical adjustment features in spending rules yields broadly similar results. Bergman and Hutchison [2015] and Combes et al. [2017] showed that fiscal policy is counter-cyclical when using fiscal rules, but there may also be non-linear effects since not all fiscal rules can reduce the procyclicality of fiscal policy when debt is high. Reuter et al. [2018] confirmed the findings of previous studies showing that fiscal rules reduce output volatility. Furthermore, not focusing the stabilization of macroeconomic variables but on the twin deficits relation, Badinger and Reuter [2015] and Afonso et al. [2018] investigated the role of fiscal rules in the relationship between countries fiscal balances and current accounts. They found a positive effect of fiscal balances on the current account, supporting the twin deficit hypothesis. However, this effect depends on the stringency of fiscal rules (budget balance or debt) in place. Following these works, it also appears that when looking at the performance of rules by studying their effects on macroeconomic variables, one can also look at any undesirable effects generated by the rules that would limit this performance. economic variables, one can also look at any undesirable effects generated by the rules that would limit this performance.

Such an investigation may be motivated and fed by some economic evidence from the last decades. Indeed, since the 1990's we saw an increase in the number of fiscal rules, as depicted in Figure 1, and the EU members are not an exception. At the same time, there has been a decline in public investment accompanied by a stagnation in GDP. Moreover, after the sovereign debt crisis, many EU countries adopted new fiscal rules, in particular debt rules, and public investment dropped to its lowest level since 1960 (according to Eurostat database). Therefore, an investigation of the side-effects of fiscal rules may constitute a major issue to tackle. Nevertheless, there are very few number of studies focusing on fiscal rules side-effects and so we try to highlight this in Chapter 3 and provide recommendations on fiscal rules implementation and design, which especially valuable since potential side-effects may mitigate fiscal rules performance. Caselli and Reynaud [2019] found that fiscal rules are correlated with lower public deficits, but the positive link disappears when endogeneity is correctly addressed. However, when considering an index of fiscal rules design, they showed that "well-designed" rules have a statistically significant impact on fiscal balance. It seems that the better fiscal rules are designed the more performant they are, making better designed fiscal rules a key target for decision-makers that adopt fiscal rules. To address this issue, it is first important to identify the negative effects of fiscal rules. This discussion on the unintended effects of fiscal rules may be of significant importance for many developing countries that have adopted rules (see figure 2), as well as for developing countries that would consider adopting them. In particular, we could hypothesise that adverse effects on investment and other economic driver variables could slow down their economic catch-up. Moreover, these considerations are also important for developed economies that face the threat of secular stagnation.

This thesis will therefore address each of the aforementioned points on the performance of the fiscal rules, namely: i) their disciplining effect on fiscal discipline; ii) the possibility to monitor them and ensuring their compliance; and iii) the identification of their side-effects that may limit their performance. But, this thesis would be incomplete without addressing the topic of measurement error: the accuracy of public finance indicators and how it can affect the performance of fiscal rules. The measurement accuracy is essential to correctly assess the performance of the rules. Errors in fiscal rules targets could lead to mistakes in the analysis of their performance.

Proposals to replace the existing rules are emerging since it is not possible to have an ideal rule (according to the fiscal rules trilemma from Debrun and Jonung [2019]), but it is possible to be satisfied with a second-best fiscal rule satisfying as many of the properties as possible. The debate is even tougher in the context of the EU with economists such as Blanchard et al. [2020] arguing in favor of replacing them by "fiscal standards" while several others proposed their own reform of the SGP to address its flaws. The famous Golden rule which targets the public budget balance excluding public net investment was put again at the center of the policymakers attention while it belongs to a 2000's debate (see for example Creel [2002] or Villieu [2003]). This rule may comply with two main characteristics of Debrun and Jonung [2019]'s fiscal rules' trilemma: it may be flexible by relaxing pressure on public productive investment and, in that sense, more easy to comply leading to greater enforceability. Nevertheless, it targets an indicator which may not be understandable nor simple to measure, leading to a lack in simplicity. Because the implementation of this rule is recommended by European Commissioner Paolo Gentiloni, we pay particular attention to this rule as a second-best proposal and discuss its measurement challenge. Indeed, such rule targets the public net investment which is subject to specific statistical measurement issues and may lead to serious mistakes. The implementation of this fiscal rule thus requires a strong statistical framework to consider it, at least, as a second-best option avoiding any fiscal rules performance mismeasurement.

## THESIS OBJECTIVES

This general introduction discussed the existence of a historic trend towards growing public debt by running public deficit. This trend may come with side-effects for economies such as macroeconomic instability, social revolts and a decrease in social welfare for future generations. To address such risks, governments have armed themselves with tools to enforce fiscal discipline, and the increase in the number of fiscal rules since the 1990's seems to reflect a willingness of governments to monitor public finance. However, the simple implementation of fiscal rules may not be enough to achieve fiscal discipline and there are thus many questions about their performance. For example, decision-makers may need to know precisely and with consistency how much the fiscal rules may reduce deficit. In that sense, there exists a need for an assessment of fiscal rules' performance on governments behavior. This question, however, only partly addresses the issue of the performance of the fiscal rules, and this thesis aims to tackle the many facets of fiscal rules performance. It considers both its beneficial and harmful (especially unwanted) outcomes, including for fiscal discipline, economic objectives and social welfare. It also addresses technical discussions to consider for the future of fiscal rules.

Qualitative approaches that followed Kopits and Symansky [1998] do not allow for an empirical assessment and quantitative conclusion on fiscal rules performance. Even if there is an exhaustive empirical literature, it is controversial due to many technical issues put forward by Heinemann et al. [2018]. In particular, the reverse causality bias mainly affects fiscal rules performance assessment studies because the fiscal rules' performance may be driven by the choice of disciplined governments to adopt fiscal rules to strengthen their commitment regarding fiscal discipline, i.e. selection bias. This discussion was raised by Wyplosz [2012] and Debrun and Kumar [2007] and Heinemann et al. [2018] empirically confirmed the existence of this bias in the majority of empirical studies focusing on fiscal rules performance. We therefore propose new empirical contributions throughout this thesis. We exploit the relative benefits of a set of approaches, relying on traditional econometric techniques in Chapter 1, before exploring the benefits of Machine Learning in chapters 2 and 3, and finally entering the statistical domain in Chapter 4. In short, this thesis offers empirical and statistical evidence on the performance of fiscal rules that appear to be connected to many economic and political fields.

The first chapter is interested in the disciplining effect of fiscal rules on public finance (i.e. the standard definition of fiscal rules performance). But, the performance also depends on goal achievement. As Kopits and Symansky [1998] defined the ideal fiscal rules as a credible rule, fiscal rules are set up to be followed. Despite Reuter [2015] dictated, an ideal fiscal rules must be credible as fiscal rules are set up to be followed. Even though Reuter [2015] has shown that even when governments do not comply with fiscal rules they are still effective (in the sense that government make effort to fulfill their objective), there is a large consensus in the literature that the credibility of fiscal rules matters. In that sense, compliance with fiscal rules also depends on the fiscal rules performance definition. For this reason, the second chapter retrieves the key determinants of the supranational fiscal rules of the EU (namely the SGP) and proposes a risk-management model to monitor them. However, excessive rigor in the rules could lead to unwanted rigidity whereas the Kopits and Symansky [1998]'s criteria called for *flexibility*. The third chapter thus studies the undesirable effects that may come with a strict application of fiscal rules. The chapter qualifies the notion of the performance of rules because they also have undesirable effects, particularly those that are detrimental to social welfare. Therefore, the notion of performance is relative and limited. Finally, by pointing out the difficulties in complying with fiscal rules as well as their side-effects, the question regarding the performance of fiscal rules arises, pushing them in the spotlight of (political) discussions regarding their future. It is therefore vital to understand how to better define them and accordingly measure their performance. Chapter 4 enters this debate by considering the emerging proposals to reform the SGP since countries like France, Spain, Italy, Greece and Portugal, that currently exceed the 100% debt-to-GDP ratio, call for an adequate reform to reflect the post-pandemic reality as well as the green transition. This chapter analyses with caution the Golden rule of public finance, and in particular the accuracy of the measurement of its target indicator, which is a key element for its performance assessment. Any error in the measurement of fiscal rules target may lead to an error in fiscal rules performance assessment. This work, conducted on the United Kingdom as a case study, is as important as it is inevitable if such a rule were to apply to all EU member states.

Here we provide a deeper summary of each chapter as follows:

**Chapter 1** of this thesis is interested in the performance of fiscal rules on fiscal discipline. Given that the EU public finance have been highly troubled by the series of economic crises, it is important to assess if the tools employed by national governments are effective to achieve fiscal discipline. This chapter proposes to answer the following question: "are national fiscal rules performant in the EU ?". It considers a rigorous definition of numerical national fiscal rules and assumes that their performance may rely on their presence-effect on fiscal discipline. The chapter conducts an empirical assessment of fiscal rules effect on fiscal discipline using a causal inference econometric approach, namely propensity-score matching, in the 28 EU economies between 2000 and 2013. It follows the first strand of literature we previously discussed which considers that the fiscal rules effectiveness is measured by their disciplinary effect on fiscal behavior. This idea is close to the first concept of Kopits and Symansky [1998] asking for well- defined fiscal rules to achieve efficient enforcement. Thus, particular attention is paid to the impact of fiscal rules on the fiscal stance trend (assessed by an indicator of public balance such as the CAPB). Many economic studies therefore focused on the impact of fiscal rules on the CAPB. The interpretation is as follows: if fiscal rules can help to increase the CAPB then they are able to ensure better fiscal discipline. This result can be found in many studies using different technical approaches (instrumental variables, system-GMM, Least Squares Dummy Variables or propensity score matching). Different conclusions are drawn in the literature, with fiscal rules appearing as non-performant in Debrun and Kumar [2007], Escolano et al. [2012] and Caselli and Reynaud [2019] where fiscal discipline is measured by CAPB, whereas many other studies take the opposite stance (see Debrun et al. [2008], Marneffe et al. [2010], Badinger and Reuter [2015] or also Combes et al. [2018]). But, Heinemann et al. [2018] addressed the issue of endogeneity bias suggesting that studies assessing fiscal rules disciplining effect may mainly suffer from reverse causality bias.

This chapter provides original thought and value-added in several ways. First, it adds to the debate on a "positive effect" versus "insignificant effect" of fiscal rules and proposes to investigate the causality of fiscal rules on fiscal discipline to provide a conclusion regarding a quantitative effect. Second, the chapter employs an empirical approach named Propensity Score Matching that addresses endogeneity problems. Third, we extend the fiscal discipline measurement from the simple CAPB to an indicator called the Global Fiscal Performance Index (GFPI) which captures a larger definition of fiscal discipline by considering public deficit, fiscal revenues, external deficit, growth rate of public debt, and growth rate of interest payments (on public debt). Indeed, CAPB may be insufficient to proxy fiscal performance and is an imperfect measure, only capturing discretionary public policy, excluding other macroeconomics variables, which could be affected by fiscal rules. Indeed, if fiscal rules entail harmful consequences on these other macroeconomic variables (while controlling discretionary policy), its effectiveness would be reduced. Finally, the analysis is conducted in the EU context where the results from the literature appear the most controversial and in opposition to one another. This chapter concludes that there is a significant positive effect of fiscal rules on fiscal discipline in the EU. Nevertheless, this positive effect depends on the type of fiscal rules as well as the economic environment. Finally, fiscal rules performance depends on several factors that also concern the definition of fiscal discipline since the alternative use of CAPB and GFPI yields different results.

**Chapter 2** investigates another definition of fiscal rules performance. It does not consider fiscal rules presence or adoption as in Chapter 1, but uses existing fiscal rules, in particular the EU supranational fiscal rules (namely SGP), and studies their compliance. It considers the following issue: "How to forecast fiscal rules compliance to strengthen fiscal rules performance monitoring?" This chapter sits alongside the literature on fiscal rules compliance such as Reuter [2019] who studied fiscal rules' compliance determinants and showed that strong fiscal rules are more complied with than others, or Larch et al. [2021] who provided empirical evidences on the positive impact of national (respectively supranational) fiscal rules on fiscal discipline. Even though fiscal rules compliance may promote fiscal discipline, Caselli and Reynaud [2019] observed a lack of compliance that raises questions on fiscal rules monitoring. Chapter 2 provides a risk-management framework to prevent compliance failures that may create distortional effects in their performance. To do so, this chapter proposes an empirical model that focuses on the EU area by conducting the analysis on the SGP. It forecasts the compliance with the SGP for the 28 EU members between 2006 and 2018. It first identifies the key determinants of SGP's compliance and then, it uses these determinants to forecast and prevent non-compliance. The model identifies 8 key features for SGP compliance that could inform policy-makers who may be interested in achieving fiscal discipline. It thus offers a framework that is easy to understand and to manage in order to monitor and ensure the credibility and performance of fiscal rules. The originality of this chapter also stems from its application of Machine Learning methodology to the issue of fiscal discipline. The Machine Learning field is becoming prominent, in particular in Finance, but such Machine Learning models have not yet been exploited in the context of forecasting fiscal discipline failures. Finally, it highlights a high forecasting score in this exercise allowing us to propose a risk-management framework for the European Union supranational fiscal rule.

**Chapter 3** addresses the limits of fiscal rules' performance by investigating their potential side-effects. It offers new perspectives on fiscal rules performance that may be mitigated by reverse-effects. While Chapter 1 provides strong evidences that the presence of fiscal rules has a positive impact on government behavior regarding fiscal performance and Chapter 2 promotes a monitoring framework to more easily prevent non-compliance with fiscal to support their credibility, it is not obvious that excessive fiscal rules enforcement is beneficial for the economy or social welfare. It is possible that enforcing fiscal discipline and imposing rigid fiscal rules' may come at a cost. This remaining issue should be tackled to cover all aspects of fiscal rules performance. That being said, there is an existing literature focusing on the impact of fiscal rules on inequalities. Combes et al. [2019] found that expenditure rules increase inequalities while budget balance rules and debt rules do not; whereas Hartwig and Strum [2019] showed, using measures of disposable income, that fiscal rules increase inequality the EU. Since the conclusions are controversial, we first try to provide a highlight in this issue. Inequalities are not the only thing that are affected by the unintended consequences of fiscal rules. If fiscal rules have stabilizing power for macroeconomics, do they operate through a cut in productive expenditure? How is the public spending composition affected by full compliance with fiscal rules?

To broaden the scope of research into the unintended effects of fiscal rules, we extend the literature by studying the following problem: "Are fiscal rules detrimental for social welfare?" This chapter therefore looks at the performance of fiscal rules through their compliance, but also puts this performance into perspective by considering different definitions of fiscal rules compliance with one strict definition of compliance as well as more flexible one. The contributions of this third chapter to the literature are numerous: First, the aforementioned literature on fiscal rules side-effects focused only on the presence or strength of fiscal rules. We extend the analysis by assessing the compliance effects. Second, focusing on social welfare allows us to study a wider field than focusing on inequalities alone. Indeed, the list of social welfare determinants may include a number of candidates such as the level of economic development, fiscal policy (Gosh and Roy [2004]) and monetary policy (Lawler [2001]), institutions (Acemoglu [2003]), the financial development (Marini [2005]), international trade (Samuelson [1938]), geography (Smith [1974])... The focus of this investigation is on OECD economies, covering 16 countries between 2004 and 2015. It carefully identifies national budget balance rules applied over a same period to provide a consistent average treatment effect. Including different types of fiscal rules and different application periods would likely come with heterogeneities which may imply too much different/opposite effects. This chapter thus proposes an empirical analysis using a causal Machine Learning approach, namely Double/Debiased Machine Learning, developed by Chernozhukov et al. [2018], to assess the impact of fiscal rules compliance on a large number of social welfare channels. This methodology is doubly relevant in this exercise as we take a two-step methodology which first identifies the key determinants for national fiscal rules' compliance and then retrieves the causal effect using a model robust against endogeneity issues. We highlight 10 major determinants for national budget balance rules compliance, including a voter preferences proxy suggesting that endogenous bias is important when assessing fiscal rules performance and controlling for voter preferences to provide robust and generalisable results. Indeed, the effects of political incentives on the fiscal discipline may depend on the fiscal preferences of the voters. If voter preferences are not considered in the specification of the model, they may introduce an endogeneity bias. The estimation of the average treatment of national budget balance rules compliance shows that fiscal rules performance may have side-effects on social welfare through a negative effect on social expenditure and an increasing effect on inequalities, in particular on the poorest groups.

The discussion of fiscal performance also relies on the accuracy of the measurement of fiscal rules targets. Each country provides economic and public finance indicators in their national accounts. The design of fiscal rules is based on public finance indicators, often expressed as a percentage of the GDP, and their performance depends on the government ability to meet these targets. Weak series for the general government sector in national accounts may imply severe errors in cross-countries comparisons, economic forecasting and nowcasting, as well as fiscal discipline assessment.

**Chapter 4** studies how sensitive the series of public net investment are to statistical functions assumed in consumption of fixed capital (CFC, or depreciation) measurement. It thus addresses the issue of the sensitivity of fiscal rules performance to statistical assumptions on capital stock measurement. These series of public net investment represent a key indicator as flexible rules, including investment-friendly rules, receive a growing attention in the current debate on fiscal rules.

In the aftermath the pandemic crisis, the needs of both investment support and fiscal sustainability co-exist with an unprecedented level of public debt in EU countries. Consequently, the debate on the possible introduction of a Golden rule of public finance in the EU context caught our attention. The post-COVID economic reality asks for new fiscal framework, in particular in monetary unions where countries were heterogeneously affected by the crisis. In the EU case, a Golden rule constitutes a proposal for the SGP reform and, while the idea may seem idyllic, the implementation still raises many constraints, not only political but also technical. Indeed, its implementation faces political issues since it requires the common acceptance from all the EU members, but this is not the discussion we ignite in this chapter. It also faces accounting and statistical challenges depending on countries national accounts harmonization, collaboration, transparency and rigor. We look to address the statistical issues to provide recommendations to properly calculate the target of the Golden rule to avoid any error in the assessment of its performance. To do so, this chapter proposes a sensitivity analysis of the United Kingdom's Golden rule to the measurement of the public CFC. It shows that the target of the Golden Rule of public finance is sensitive to statistical assumptions and parameters utilised in capital stock measurement from which CFC is derived. In particular, the age-profile and the depreciation rate appear as major determinants for data series of public CFC and thus public net investment. Consequently, this chapter argues that if a Golden rule were chosen as a new European fiscal rule, and if it was considered a second-best option, a strong accounting strategy across European countries would be vital. This chapter concludes that it is not impossible to propose a rule which limits certain perverse effects of overly strictly applied fiscal rules while maintaining their disciplining nature, but all this is possible under correct application and alignment of statistical conditions. Otherwise, weak calculations of the target could lead to errors in judgment and result in performance deviations. Lessons related to the sensitivity of the measurement of the public CFC, which directly impacts the measurement of public net investment (net of depreciation), is also crucial for the measurement of GDP. Indeed, the value added of the public sector (or general government to be more precise) contributes to the measure of total GDP and the CFC enters in the composition of GDP. Affecting the CFC therefore impacts the GDP, bringing with it consequences for nowcasting, forecasting, the measurement of potential output and the output gap, as well as for all variables expressed as a percentage of GDP. Fiscal rules targets are mainly expressed as a percentage of GDP and may exclude cyclical components. It is clear that the Golden rule isn't the only fiscal rule affected by the accuracy of the methodology to compute net capital stock and retrieve CFC. The value-added of this analysis is therefore profound with its conclusions concerning a wide range of national accounts indicators.

# Bibliography


D. Acemoglu. Why not a political coase theorem? Social conflict, commitment and politics. *Journal of Comparative Economic Studies*, pages 620–652, 2003. 16

A. Afonso, F. Huart, J. Tovar Jalles, and P. Stanek. Fiscal Rules and Twin Deficits: The Link between Fiscal and External Balances. *REM Working Paper 031-2018*, 2018. 9

H. Badinger and W. H. Reuter. Measurement of fiscal rules: Introducing the application of partially ordered set (POSET) theory. *Journal of Macroeconomics*, 43:108–123, 2015. 8, 9, 14

R. Barro. Are government bonds net wealth ? *Journal of Political Economy*, page 1095–1117, 1974. 3

M. Bergman, M.M. Hutchison, and S.E. Hougaard Jensen. Promoting sustainable public finances in the European Union: The role of fiscal rules and government efficiency. *European Journal of Political Economy*, 44:1–19, 2016. 8

U. M. Bergman and M. Hutchison. Economic stabilization in the post-crisis world:Are fiscal rules the answer? *Journal of International Money and Finance*, 52:82–101, 2015. 9

O. Blanchard, A. Leandro, and J. Zettelmeyer. Redesigning the EU Fiscal Rules: From Rules to Standards. *Mimeo, contribution presented to the 72nd Economic Policy Panel Meeting*, 2020. 11

H. Bohn and R.P. Inman. Balanced Budget Rules and Public Deficits: Evidence from the U.S. States. NBER Working Paper No. 5533, 1996. 7, 9

J. B. Burbidge. Government Debt in an Overlapping-Generations Model with Bequests and Gifts. *The American Economic Review*, pages 222–227, 1983. 2

M. Buti, S. Eijffinger, and D. Franco. Revisiting the Stability and Growth Pact: grand design or internal adjustment?). *European Economy-Economic Papers n180*, 2003. 7

F. Caselli and J. Reynaud. Do Fiscal Rules Cause Better Fiscal Balances? A New Instrumental Variable Strategy. *IMF Working Paper No. WP/19/49*, 2019. 8, 10, 14, 15

V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, Ch. Hansen, W. Newey, and J. Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 2018. 16

J-L. Combes, A. Minea, and M. Sow. Is fiscal policy always counter- (pro-) cyclical? The role of public debt and fiscal rules. *Economic Modelling*, 65:138–146, 2017. 9

J-L. Combes, A. Minea., P.N. Sawadogo, and C. Vinturis. Can fiscal rules curb income inequality? Evidence from developing countries. *CERDI Etudes et Documents No. 25*, 2019. 16

J-L. Combes, X. Debrun, A. Minea, and R. Tapsoba. Inflation Targeting, Fiscal Rules and the Policy Mix: Cross-effects and Interactions. *The Economic Journal*, 128:2755–2784, 2018. 14

European Commission. Analysis Fiscal framework reformes pfr 2010. In *Public finances in EMU*, chapter Chap4 Fiscal rules, independent institutions and medium-term budgetary frameworks. 2010. 3

J. Creel. Asymétries budgétaires dans la zone euro, un essai de modélisation du Pacte de stabilité. *Revue française d'économie*, 16:91–127, 2002. 7, 11

X. Debrun and L. Jonung. Under threat: Rules-based fiscal policy and how to preserve it. *European Journal of Political Economy*, 57:142–157, 2019. 8, 11

X. Debrun and M.S. Kumar. The Discipline-Enhancing Role of Fiscal Institutions: Theory and Empirical Evidence. *IMF Working Paper No. 07/171*, 2007. 7, 12, 14

X. Debrun, L. Moulin, A. Turrini, J. Ayuso i casals, and M. Kumar. Tied to the Mast? National Fiscal Rules in the European Union. *Economic Policy*, 23:297–362, 2008. 8, 14

B. Egert. Public debt, economic growth and nonlinear effects : myth or reality ? *Journal of Macroeconomics*, pages 226–238, 2017. 3

J. Escolano, L. Eyraud, J. Sarnes M. Moreno Badia, and A. Tuladhar. Fiscal Performance, Institutional Design and Decentralization in European Union Countries. *IMF Working Paper No. 12/45*, 2012. 8, 14

D. Foremny. Sub-national deficits in European Countries: The impact of fiscal rules and tax autonomy. *European Journal of Political Economy*, 34:86–110, 2014. 8

S. Gosh and U. Roy. Fiscal policy, long-run growth, and welfare in a stock-flow model of public goods. *Canadian Journal of Economics/Revue canadienne d'économique*, pages 742–756, 2004. 16

M. Guerguil, P. Mandon, and R. Tapsoba. Flexible fiscal rules and countercyclical fiscal policy. *Journal of Macroeconomics*, 52:189–220, 2017. 9

J. Hartwig and J.E. Strum. Do fiscal rules breed inequality? First evidence for the EU. *Economic Bulletin*, 2019. 16

F. Heinemann, M-D. Moessinger, and M. Yeter. Do fiscal rules constrain fiscal policy? A meta-regression-analysis. *European Journal of Political Economy*, 51:69–92, 2018. 12, 14

G. Kopits and S.A. Symansky. Fiscal Policy Rules . *Occasional Paper No. 162*, 1998. 6, 7, 8, 12, 13

F. Kydland and E. Prescott. Rules Rather Than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy*, 85:473–491, 1977. 3

M. Larch and S. Santacroce. Numerical compliance with EU fiscal rules: The compliance database of the Secretariat of the European Fiscal Board (Database). 2020. 8

M. Larch, E. Orseau, and W. van der Wielen. Do EU Fiscal Rules Support or Hinder Counter-Cyclical Fiscal Policy?). *Journal of International Money and Finance*, 2021. 15

P. Lawler. Monetary policy, central bank objectives, and social welfare with strategic wage setting. *Oxford Economic Papers*, page 94–113, 2001. 16

F. Marini. Banks, financial markets, and Social Welfare. *Journal of Banking and finance*, pages 2557–2575, 2005. 16

W. Marneffe, B. Van Aarle, W. Van Der Wielen, and L. Vereeck. The Impact of Fiscal Rules on Public Finances: Theory and Empirical Evidence for the Euro Area. *CESifo Working Paper No. 3303*, 2010. 8, 14

F. Modigliani. Long-run implications of alternative fiscal policies and the burden of the national debt. *Economic Journal*, 71:730–755, 1961. 3

R. Mundell. Capital mobility and stabilization policy under fixed and flexible exchange rate. *Canadian Journal of Economics*, 29:475–485, 1963. 2

M. Wafula Nandelenga and M. J. Ellyne. Fiscal Rules and the Compliance debate: Why do Countries adopt Rules and fail to comply. *ERSA Working Paper No 815*, 2020. 8

E. Perrot. La dette publique. *Études 2010/2 (Tome 412)*, pages 177–188, 2010. 1

T. Persson. Deficits and intergenerational welfare in open economies. *Journal of International Economics*, 19:473–491, 1985. 2

A. C. Pigou. Macmillan, london, 2nd edition. *A Study in Public Finance*, pages 185–203, 1929. 2

C. Reinhart and K. Rogoff. Growth in a time of debt. *American Economic review*, pages 573–578, 2010. 3

W. H. Reuter, O. Tkačevs, and K. Vilerts. On the design of stabilising fiscal rules. *Arbeitspapier No. 11/2018*, 2018. 9

W.H. Reuter. National numerical fiscal rules: not complied but still effective? *European Journal of Economy*, 39:67–81, 2015. 12

W.H. Reuter. When and why do countries break their national fiscal rules? *European Journal of Political Economy*, 57:125–141, 2019. 8, 15

A. Sacchi and S. Salotti. The impact of national fiscal rules on the stabilisation function of fiscal policy. *European Journal of Political Economy*, 37:1–20, 2015. 9

P. Samuelson. Welfare Economics and International Trade. *The American Economic Review*, pages 261–266, 1938. 16

A. Schaechter, T. Kinda, and N. Budinaand A. Weber. Fiscal Rules in Response to the Crisis—Towards the 'Next-Generation' Rules. A New Dataset. *IMF Fiscal Affairs Department*, 2012. 4

B. Schwengler. Les règles budgétaires, une alchimie délicate. *Revue d'économie française*, XXVII:226, 2012. 3

D. M. Smith. Who gets what Where, and How: A welfare focus for human geography. *Geography*, pages 289–297, 1974. 16

P. Villieu. Pacte de stabilité, crédibilité du policy mix et coordination des politiques budgétaires en union monétaire. *Revue économique*, 54, 2003. 11

C. Wyplosz. Fiscal Rules: Theorical issues and historical experience. *NBER Working Paper No. 17884*, 2012. 2, 12

# Chapter 1

# National fiscal rules and fiscal discipline in the European Union

"The strongest case for fiscal rules is rooted in political economy. In a democratic society, rules are necessary to restrain politically rational policymarkers who conduct discretionary policies with a deficit bias when facing an electorate that fails to understand, or is indifferent to, the intertemporal budget constraint" Kopits [2001]

*A version of this chapter, co-authored with A. Barbier-Gauchard (University of Strasbourg, France) and A. Minea (University of Clermont-Auvergne, France), is published in Applied Economics.*

## 1.1 Introduction

The Global Financial Crisis followed by the sovereign debt crisis has heightened attention to the trend of increasing public spending since the late 1990s, thereby reviving considerations on fiscal rules. The birth of the Economic and Monetary Union, and its link with the Optimal Currency Areas Theory of Mundell, has largely raised the question of the desirable level of convergence discussed by Jacques Delors. The prospects of a single currency made necessary the introduction of a fiscal framework in the EU. The Stability and Growth Pact was born in 1996 to strengthen the monitoring and coordination of national fiscal policies. The desire to continue monitoring public finance after eurozone accession can constitute a first argument for the creation of the Stability and Growth Pact. From a broad perspective, fiscal rules are intended to discipline governments and allow for more confidence between governments, financial markets and citizens. In the eurozone context, this issue is even more crucial since it also concerns the financial stability of the monetary union: a member country in financial difficulty could lead to destabilizing effects of the monetary union as a whole (i.e. a domino effect). National governments have gradually adopted their own national fiscal rules, in addition to the supranational fiscal rule of the SGP entered into force with the birth of the EMU in 1999. Nevertheless, their strict application is rarely satisfied by the eurozone members, and the European Union (EU) members in general. The sanctions provided by the Pact were never applied, and this is why it seems to have lost all the credibility necessary for the constitution of an effective fiscal rule defined by Kopits and Symansky [1998]. Consequently, it may be legitimate to ask: are fiscal rules performant? In this context, it seems essential to assess the effectiveness of fiscal rules by studying their effect on fiscal discipline. Chapter 1 tackles this problem by studying the effect of national fiscal rules presence on fiscal discipline, in the EU between 2000 and 2013.

The study of the relationship between fiscal rules and fiscal discipline became prominent,[1] to the point where the number of existing studies was sufficiently high to fuel the recent meta-analysis of Heinemann et al. [2018]. One of the most interesting conclusion of their analysis is that—while overall fiscal rules provide more fiscal discipline by reduc-

[1] Aside from fiscal discipline, other studies analyze the impact of fiscal rules on various aspects of the fiscal policy, and in particular fiscal policy discretion (see e.g. Badinger [2009]) or cyclicality (see e.g. Bergman and Hutchison [2015] or Combes et al. [2017]), or fiscal consolidations (see e.g. Guichard et al. [2007] and Bamba et al. [2020]).

ing deficits—the opposite may arise for Euro area countries: fiscal rules seem to be associated with *increased* deficits. Although this striking finding may be consistent with the fiscal imbalances experienced by some EU countries during the recent sovereign debt crisis, it calls for a careful reassessment. Consequently, the aim of this chapter is to analyze if national fiscal rules can indeed shape fiscal behaviors in the EU, towards achieving higher fiscal discipline.

Compared with the existing literature on fiscal rules and fiscal discipline, this chapter is designed as follows. First, similar to Debrun et al. [2008], we focus exclusively on EU countries, and—in particular—we do not mix them with developing countries as in Combes et al. [2018]. Second, we take at heart to incorporate the suggestions of Heinemann et al. [2018], and particularly the fact that the favorable impact of fiscal rules on fiscal discipline is weakened if the possible endogeneity is not controlled for. While recent studies on the EU countries draw upon regression-based methods, including IV (Foremny [2014]), LSDV (Reuter [2015]) or system-GMM (Bergman et al. [2016]), we follow the work of Tapsoba [2012] performed on developing countries, and draw upon quasi-experimental methods—namely, propensity score matching. As such, we account for the issue of self-selection, i.e. the fact that governments may adopt fiscal rules because of a bad cyclically adjusted primary balance. Third, as illustrated by Heinemann et al. [2018], fiscal rules affect fiscal discipline in various ways depending on the measure of the former (e.g. deficit, debt, expenditure, or revenue) and of the latter. We first used the popular measure of fiscal discipline used in the existing literature—namely the cyclically-adjusted primary balance (CAPB), see Tapsoba [2012]— computing using the Hodrick-Prescott's filter. In fact, the Hodrick Prescott filter makes possible to extract the cycle from an economic series, and constitutes an approximation of a band-pass filter which eliminates the highest and lowest frequencies (Ahamada and Jolivaldt). Because this approach is imperfect, we also use a trigonometric filter based on Fourier approach. Finally, using these cycle extraction filters to isolate the primary structural public balance corresponds to a residual approach. Indeed, the method responds well to the definition of the public balance after extraction of cyclical elements. After subtracting the cycle from the public balance, the structural public balance should be the residual of the specification. However, the volatility of the residuals of econometric estimates poses problems of validity of the results. This is why we also propose to use the estimates of the primary structural public balance provided by the IMF and using the production function approach, capturing the real contributions of growth to the public balance. Despite the use of these different measures of CAPB help in making our results robust, it only captures a little part of fiscal discipline and we thus draw upon an original measure of fiscal discipline, namely a Global Fiscal Performance Index (GFPI). We compute this index by a two-stage approach, with the aim of going beyond single-variable measures (such as the CAPB) in order to capture the various facets of the wide concept of fiscal discipline. In particular, this original measure of fiscal discipline has the advantage of not being sensitive to the estimation of a cyclical component, since it does not require the use of filtering (or of assumptions on elasticities, as in production function approach). Fourth, we pay attention to the selection of fiscal rules. Following Debrun et al. [2008] and Reuter [2015], we drop from our sample the rules that are mostly related to the Medium Term Budgetary Framework (MTBF). This is because—as indicated on the European Commission (EC) website devoted to them—the MTBFs display some notable differences with respect to the traditional definition of fiscal rules of Kopits and Symansky [1998] (namely, "*a sustainable constraint on fiscal policy under the form of a numerical target on a key aggregate of public finances*"); such difference are related to, for example, the considered horizon—usually "*beyond the annual budgetary calendar*", and the form of commitment—usually "*a weaker form of commitment than a pure rule incorporating binding targets*" (see the EC website). By doing so, we improve the homogeneity of our measure of fiscal rules. Finally, Heinemann et al. [2018] suggest that the effect of fiscal rules on fiscal discipline may differ with respect to the characteristics of the study. We explore three sources that may affect the impact of fiscal rules on fiscal discipline, namely: the method used; the type of fiscal rule; and the countries' structural characteristics.

Chapter 1 provides the following results:

(1) EU countries that present fiscal rules significantly improve their fiscal discipline—measured by the CAPB, computed using three alternative measures of the output gap—with respect to comparable EU countries without fiscal rules. Contributing to the debate on the effect of fiscal rules on the CAPB (for example, Debrun and Kumar [2007] reveal the lack of a significant response of the CAPB when fiscal rules are instrumented, which is consistent with the conclusions of Escolano et al. [2012] on the group of EU15 countries, while the response of CAPB is significant and positive in Debrun et al. [2008] and Marneffe et al. [2010]),[2] our study reveals that fiscal rules have a favorable effect on the CAPB in our

[2]Such a debate is equally at work when differentiating countries depending on their level of economic development: the response of the CAPB is not significant in the sample of 49 advanced and emerging market

treatment effect analysis controlling for reverse causality.

Capitalizing on this finding, we show that the presence of fiscal rules enforces fiscal discipline captured by our novel measure, namely the GFPI. These findings—supported by various tests for the quality of the matching—are robust across different matching methods, when using an alternative estimator (namely, the doubly robust inverse-probability-weighted regression adjustment, IPWRA), when using an alternative computation for the GFPI, and when further increasing the vector of control variables or altering the sample.

(2) There are important differences in the effect of the various types of fiscal rules—namely, Balance Budget Rules (BBR), Expenditure Rules (ER), and Debt Rules (DR)—on fiscal discipline. Specifically, while BBR (ER) significantly improve (leave statistically unchanged) the CAPB consistent with the existing literature, contrary to previous studies our estimations do not support a significant impact of DR on the CAPB. Moreover, while BBR and DR conserve their respective effect on fiscal discipline measured by the CAPB, we show that ER significantly improve the GFPI with a magnitude larger than that of fiscal rules altogether. Subsequent estimations performed using the variables that compose the GFPI confirm that the effect of the various types of fiscal rules can indeed differ—both in significance and magnitude—with the measures of fiscal discipline. In particular, while both BBR and ER significantly reduce the public deficit and the growth of public debt, and ER and DR lower the growth rate of public expenditure, only BBR (ER) significantly decrease the growth of interest rate (the external deficit).

(3) The effect of fiscal rules on fiscal discipline is subject to important heterogeneities, related to macroeconomic factors, political factors, and factors associated with the fiscal rules themselves. Three types of results emerge when comparing the influence of these factors on the effect of fiscal rules on the CAPB and the GFPI: some variables, such as the public debt ratio, reduce the favorable effect of fiscal rules on both measures of fiscal discipline; other variables, such as the real GDP per capita, do not exert a significant impact on the effect of fiscal rules on fiscal discipline irrespective of its measure; finally, some variables significantly affect only the CAPB—for example, bad times (the number of years covered by the rule) weaken (foster) the favorable effect of fiscal rules—, or only the GFPI—for example, the presence of the SGP fosters the favorable effect of the fiscal rules.

The rest of the chapter is structured as follows. Section 2 briefly reviews the related

economies for Cevik and Teksoz [2014], but significant and positive for Tapsoba [2012] who considers 74 developing countries.

literature on fiscal rules and fiscal discipline, Section 3 describes the data by insisting on our novel measure of fiscal discipline, Section 4 presents the methodology, Section 5 reports the benchmark results, Section 6 assesses their robustness, Section 7 investigates the presence of heterogeneity in the effect of fiscal rules on fiscal discipline, and Section 8 concludes the chapter.

## 1.2 Literature review on fiscal rules and fiscal discipline

### 1.2.1 Fiscal rules as a policy to promote fiscal discipline: Theory

Rooted in the late the 1980s, fiscal discipline in the European and Monetary Union (EMU) is at the core of the European construction. From a broad perspective, fiscal discipline is related to the ability of a government to maintain sustainable public finance. Several tools could reach this objective namely fiscal consolidation programs (see Bamba et al. [2020]) or fiscal rules. Irrespective of the complex issue of its measurement (see the next section), fiscal discipline in EMU is aimed to be ensured—as previously emphasized—by the supranational fiscal rules of the SGP and national fiscal rules.

Resting upon fiscal rules to ensure fiscal discipline is supported by several arguments developed by Wyplosz [2013]. Perhaps the most popular of them is the political economy viewpoint on "the political bias for public deficit", according to which public policymakers' behavior may lead to public deficits above those that would correspond to an optimal fiscal policy.[3] Such a behavior can arise mainly as the result of (*i*) the well-known "tragedy of the commons" related to the common pool problem that may generate free-riding behaviors (Velasco [2000]), and (*ii*) the policymakers' short-term horizon due to their particular interest for the upcoming elections (Alesina and Tabellini [1990]) that may translate into time-inconsistent government policies (Persson et al. [2006]). From this perspective, fiscal rules may improve the temporal coherence of fiscal policies by disciplining governments and building confidence on the financial markets, which is expected to ultimately limit excessive debt financing of fiscal policy and improve fiscal discipline.

[3] For example, excessive public deficits may be a threaten for a monetary and economic union as a whole, as they exacerbate the systemic risk and yield domino effects (see e.g. Camdessus [1999] or Kumar and Ter-Minassian [2007]); Krogstrup and Wyplosz [2010] discuss the issue of supranational deficit ceilings.

### 1.2.2 Measuring the fiscal rules performance on fiscal discipline

A large literature investigates the impact of fiscal rules on fiscal discipline, usually approached by an indicator of the fiscal balance. In an early contribution, Bohn and Inman [1996] show that government balance requirements significantly affected U.S. states' general fund surplus during the 1970-1991 period. Such a favorable effect of fiscal rules on fiscal discipline is equally supported by e.g. Debrun et al. [2008], Marneffe et al. [2010], who measure fiscal discipline by the CAPB.[4]

However, the existing literature identified at least three possible sources that may weaken this favorable effect of fiscal rules on fiscal discipline. A first source—of methodological order—is related to the issue of endogeneity. Debrun and Kumar [2007] reveal the lack of a significant response of the CAPB when fiscal rules are instrumented, consistent with the conclusions of Escolano et al. [2012] on the group of EU15 countries and of the more recent analysis of Caselli and Reynaud [2019] performed on a large panel of 142 countries.

A second source—of measurement order—is related to the way fiscal discipline is captured. For example, Debrun et al. [2008] found a significant effect of fiscal rules on the CAPB, a result extended by Afonso and Hauptmeier [2009] to the primary balance. However, while confirming the favorable effect of fiscal rules on e.g. the primary and the overall fiscal balance, Marneffe et al. [2010] equally revealed that fiscal rules do not significantly affect government revenues, the cyclical fiscal balance, or the change in the structural primary balance.

Finally, a third source is related to heterogeneity and conditionality. Regarding the former, the literature points out to an unequal impact of the various types of fiscal rules on fiscal discipline. While some studies focus on a particular form of fiscal rules (see e.g. Holm-Hadulla et al. [2012] who emphasize a favorable effect of expenditure rules on government expenditure), others compare the impact of different types of fiscal rules. In the case of developing countries, Tapsoba [2012] shows that, contrary to balance budget and expenditure rules that significantly improve the CAPB, debt rules have no significant effect. However, focusing on the EU countries, Bergman et al. [2016] conclude that balance budget rules are more effective than expenditure and debt rules to increase the CAPB, while debt rules significantly improve the CAPB by themselves—and not only when combined with balance budget rules as in Debrun et al. [2008]. Regarding the latter, several studies (e.g. Tapsoba [2012], or Combes et al. [2018]) outline that the impact of fiscal rules

[4] Comparable conclusions arise from the analysis Foremny [2014] performed on subnational fiscal rules.

on fiscal discipline may differ in various environments, be them fiscal, monetary, or institutional.

By taking stock of these findings, our study aims at revisiting the complex relationship between fiscal rules and fiscal discipline.

## 1.3 Data

This section is devoted to the presentation of our main variables, namely measures of fiscal discipline (the dependent variable) and fiscal rules (the main independent variable).

### 1.3.1 The measure of fiscal discipline

As discussed by Minea and Tapsoba [2014] and Hallerberg et al. [2009], fiscal discipline is a complex concept that can be approached in several ways. Most of the studies devoted to fiscal discipline usually capture it using a single variable providing information about a fiscal aggregate. The literature on fiscal discipline and fiscal rules makes no exception: in their meta-analysis, Heinemann et al. [2018] consider studies that measure fiscal discipline by fiscal deficit, debt, expenditure, or revenue. Since our goal is not to be exhaustive about the different single-variable measures of fiscal discipline, we first focus on the popular Cyclically Adjusted Primary Balance (CAPB). However, subsequently, we go beyond the CAPB, and build an original measure of fiscal discipline designed to better seize its complexity.

**The Cyclically-Adjusted Primary Balance (CAPB) as the traditional measure of fiscal discipline**

The CAPB was used to measure fiscal discipline in the analysis of Tapsoba [2012] devoted to developing countries. Since it is not directly observable, we estimate it using the residual approach of Fatás and Mihov [2003, 2006]

$$\mathrm{PBB}_{i,t} = \alpha + \beta \mathrm{PBB}_{i,t-1} + \gamma \mathrm{GAP}_{i,t} + \varphi \mathrm{W}_{i,t} + \eta_t + \epsilon_{i,t}, \tag{1.1}$$

with $\mathrm{PBB}_{i,t}$ the primary budget balance. To properly isolate the CAPB through the error term $\epsilon_{i,t}$, i.e. the residual of the PBB after extracting the cyclical elements, we perform several corrections: (i) to avoid an endogeneity problem, the output gap ($\mathrm{GAP}_{i,t}$)—computed using the popular Hodrick-Prescott (HP) filter with a smoothing parameter set at 100 given the use of yearly data—is instrumented by its own lagged value;[5] (ii) we follow Turner [2006], and control in $W_{i,t}$ by inflation and the terms of trade; and (iii) according to Villafuerte and Lopez-Murphy [2010], we account for the price of raw materials through the time fixed effects $\eta_t$.

Despite its popularity, the CAPB has at least two shortcomings. First, the CAPB is contingent to the method used to calculate the output gap, which is an unobserved variable. Given that there is no consensus on the best method to compute the output gap (see e.g. Andersen [2013]), we use—in addition to the popular Hodrick-Prescott (HP) filter—two alternative ways for the computation of the output gap (and therefore of the CAPB), namely a trigonometric filter and the production function approach. Second, the CAPB focuses exclusively on the discretionary fiscal policy. Precisely to cope with this shortcoming, we develop a novel measure of fiscal discipline.

**Fiscal discipline measured by the Global Fiscal Performance Index (GFPI)**

As underlined previously, the CAPB appears to be an insufficient indicator to assess the effectiveness of fiscal rules. Indeed, by considering only the discretionary fiscal policy, the CAPB is unable to capture the potential impact that fiscal rules may have on other macroeconomic aggregates, such as the level of public debt, the interest rates on the public debt, the variations of public revenues/expenditures or even the external balance (due in particular to the potential presence of twin deficits). For example, in the Macroeconomic Imbalance Procedure (MIP) Scoreboard, the EC is monitoring a broad set of macroeconomic aggregates to capture the risks of macroeconomic imbalances, including e.g. the public deficit or the external balance (see the Alert Mechanism Report 2020 of the European Commission [2019]). Consequently, to seize the multiple facets of fiscal discipline, we build an aggregated measure—the *Global Fiscal Performance Index* (GFPI).

Our approach to build the GFPI is inspired by the work of Mohanty and Mishra [2016], who—capitalizing on the methodology of the United Nations Development Programme (UNDP)—use five principal indices (obtained from several primary indicators) to compute the index for seventeen Indian states. However, compared with Mohanty and Mishra [2016], our methodology differs on two grounds. First, to account for the specificities of

[5]We instrument the output gap with its own lagged value; since in this equation we control for country fixed effects, we use a system-GMM estimator that appropriately deals with the dynamic panel bias of Nickell [1981].

the EU—and particularly the monetary union features and its role in international trade—we consider an external position indicator that pays attention to twin deficits (see the discussion in e.g. Badinger et al. [2017] and Afonso et al. [2018]). Second, to avoid compensation between indices, our methodology consists of two stages that involve popular methods for the construction of composite indices, detailed in the following.

In Stage 1, using five primary indicators of public finance (see Table 1), namely, public deficit, fiscal revenues (considered with a negative sign for consistency with the other indicators), the external deficit, the growth rate of public debt, and the growth rate of interests (on public debt), we obtain four secondary indices by standardization,[6] that reflect respectively the risk of high deficit, the risk of insufficient collected revenues, the risk of external imbalance, and the risk of unsustainability. In particular, the latter index contains information from the latter two primary indicators (the growth rate of public debt and the growth rate of interests on public debt), and is computed using the Mazziotta-Pareto approach.[7]

[6]According to the Competence Centre on Composite Indicators and Scoreboards (COIN) of the European Commission, "*the normalized indicator value for a country is calculated as the ratio of the difference between the raw indicator value and the average divided by the standard deviation.*"

[7]The computation is as follows: assume $X = \{x_{ij}\}$ is a matrix with n units (rows) and m indicators (columns), $M_{x_j}$ is the mean for the indicator j, and $St_{x_j}$ its standard deviation; then, the normalized matrix $Z = \{z_{ij}\}$ is calculated as $z_{ij} = 100 \pm \frac{x_{ij} - M_{x_j}}{St_{x_j}} 10$, where $\pm$ give the polarity of the indicator j. With $M_{z_i}$ and $St_{z_i}$ the mean and the standard deviation of the standardized values for the unit i, respectively, the Mazziotta-Pareto Index (MPI) can be written as $MPI_i^{\pm} = M_{z_i} \pm St_{z_i} cv_i$, with $cv_i = St_{z_i}/M_{z_i}$ the coefficient of variation for unit i. In our analysis, the higher the index, the higher the risk of unsustainability.

Table 1.1: **GFPI primary indicators**

| *Primary Indicator* | *Variable* | *Source* | *Level 2 index construction* |
|---|---|---|---|
| Fiscal Revenues | Total revenues of public administrations (including taxes) in % of GDP | Eurostat: Main aggregates of general government, including revenues and expenditures | Variation between $t$ and $t-1$ in fiscal revenues considered with a negative sign for consistency with the other indicators + standardization procedure. |
| Public Deficit | Total fiscal balance in % of GDP | IMF Fiscal Indicators: Net lending/borrowing (also referred as overall balance) in % of GDP) | Deficit in total fiscal balance (negative sign in the total fiscal balance traduces a fiscal deficit which implies a positive level 2 index) + standardization procedure. |
| External deficit | Net External Position (NEP) in % of GDP | Eurostat: Macroeconomic Imbalance Procedure Indicators/External Debt/Net External Position. The indicator is a subset of the NEP that excludes equity-related components, such as holdings and participations in foreign direct investment, and intra-instrument debt. The indicator is defined as NEP minus net direct investment minus net portfolio investments. | Deficit in NEP (negative sign in the NEP traduces an external external deficit which implies a positive level 2 index) + standardization procedure. |
| Public debt growth rate | Debt on GDP ratio (in % of GDP) | IMF Historical Database | Growth rate of both indicators come from authors' calculations |
| Interest growth rates (on public debt) | Interest payments (in % of GDP) Authors' calculations to obtain the indicator in % of GDP. | World Bank Development Indicators Interest payments include interest payments on government debt–including long-term bonds, long-term loans, and other debt instruments –to domestic and foreign residents | and are aggregated with Marriota-Pareto index aggregating approach + standardization procedure. |

In Stage 2, we aggregate the four secondary indices using the "*Mean-Min Function*" into the MMF index, defined as $\mathrm{MMF}_i = \mathrm{M}_{z_i} - \alpha\left(\sqrt{(\mathrm{M}_{z_i} - \min_j\{z_{ij}\})^2 + \beta^2} - \beta\right)$, with $z_{ij}$ the element of the matrix Z of our normalized indexes, $\mathrm{M}_{z_i}$ the average of the standardized values, $0 \leq \alpha \leq 1$ the intensity of penalty for imbalances, and $\beta \geq 0$ the intensity of the complementarity between the indicators. We checked beforehand if these variables are not too correlated, to avoid the risk of counting some effects several times when aggregating them (Appendix 3 reports the correlation matrix). This index is independent of the choice of the indicator normalization procedure, and since $\alpha \neq 0$ ($\alpha = 0$ corresponds to the arithmetic mean) it avoids compensation in order to capture the effect of each indicator. After taking the opposite sign of this index and normalizing the values, we obtain

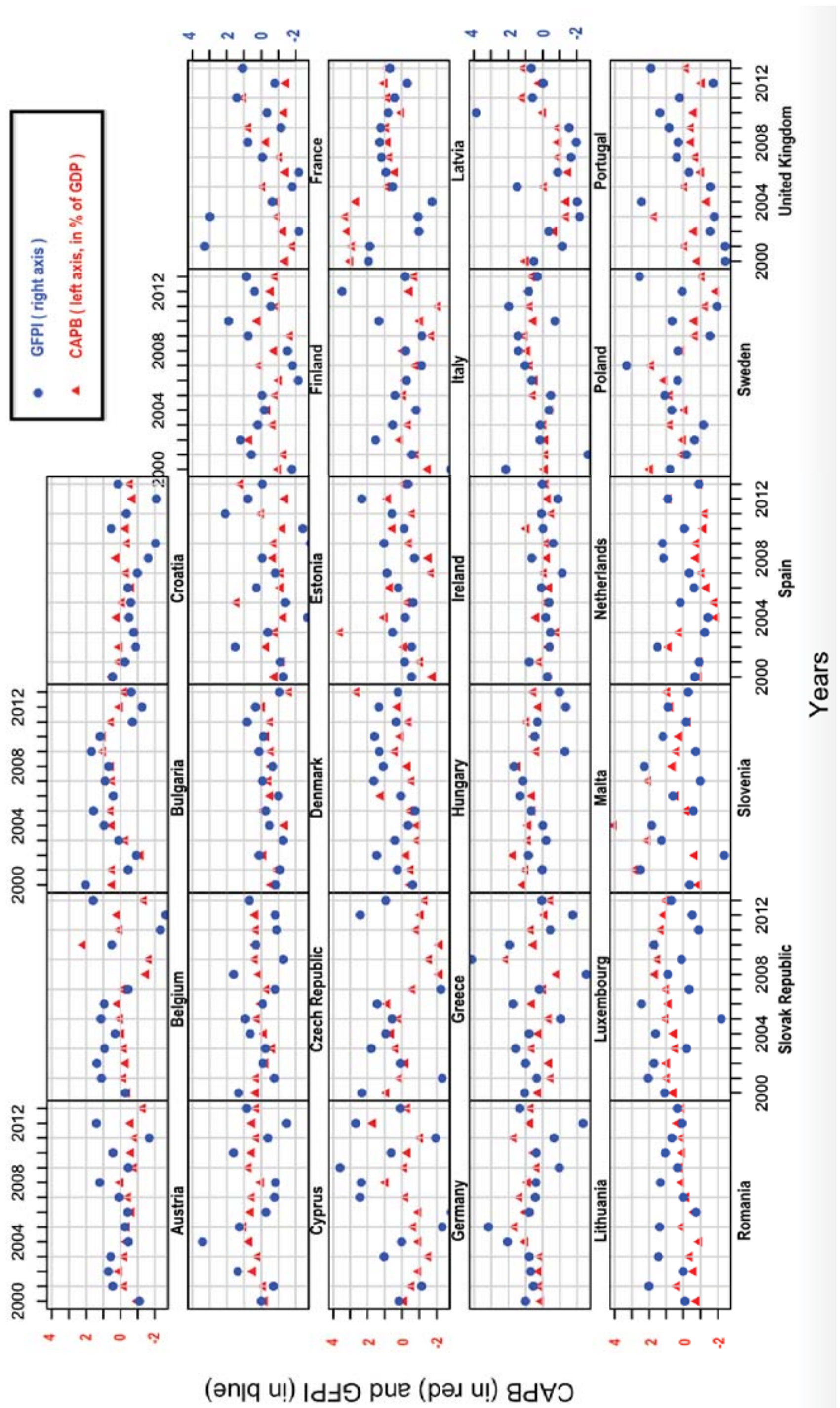



Figure 1.1: **CAPB and GFPI in the EU countries in our sample**

our GFPI index; comprised between -2 and 4, a higher value of this yearly-frequency index signals better fiscal performance (see Appendix 10 for an illustration of the two stages).

Figure 1 plots the GFPI index and the traditional measure of fiscal discipline (the CAPB) for the EU countries in our sample. A simple visual inspection reveals the differences between the two measures of fiscal discipline. In particular, the GFPI seems more volatile than the CAPB for most countries, especially around crisis periods (the *dotcom* bubble in the early 2000s and the Great Recession in the late 2000s). This may be related to the fact that, while the CAPB is smoothed out of the cycle, the GFPI—by embedding information for several variables, including e.g. public debt or interest rates—reproducing the dynamics of these various variables that may be particularly important around crisis times.

Finally, we back up these simple observations with a more formal statistical comparison between the CAPB and the GFPI, based on the Least Absolute Shrinkage and Selection Operator (LASSO) selection procedure developed by Tibshirani [1996]. LASSO is a Machine Learning tool that performs regularization and feature selection, by applying a regularization process where the coefficients of the less important predictors are penalized and shrunk to zero (for further details, see Appendix 14). Our exercise consists of forecasting the change in public debt during the period 2001-2014 using a wide range of potential predictors, namely: CAPB, GFPI, inflation, trade openness, government stability, external deficit, financial liabilities, gross fixed capital formation, and real GDP per capita growth. Among all these variables, three are found to matter for forecasting the change in debt, namely: CAPB, GFPI and the real GDP per capita growth. In addition, we can estimate a coefficient that provides a measure of the explanatory power of a variable for the change in debt, which equals to 0.169 for the CAPB and 0.867 for the GFPI. These findings suggest that the GFPI outperforms the CAPB in explaining the change in debt.[8]

### 1.3.2 Fiscal rules

During the last decades the number of fiscal rules increased in the EU. Compared to only two countries in 1990 (Germany adopted a balanced budget rule in 1969 and a public expenditure rule in 1982, and Luxembourg adopted a debt and a public expenditure rule in 1990), in 2015 all EU countries had at least one national fiscal rule. However, to mitigate the influence of the numerous fiscal rules enacted in response to the recent sovereign debt

[8]We are indebted to an anonymous referee for suggesting that we perform this analysis.

crisis, we restrain our analysis until 2013.[9] Nevertheless, we are still capturing flexible fiscal rules (see Guerguil et al. [2017]), including e.g. rules that favor investment, rules that include escape clauses, and rules with cyclically-adjusted goals; therefore, we checked that the selected rules are compatible with the supranational framework and also with the flexibility necessary for cyclical adjustment. Consequently, after equally excluding the MTBFs, out of the twenty-eight countries in our sample, twenty countries had at least one national numerical fiscal rule and eight countries did not adopt national fiscal rules by 2013 (see Appendix 1 for the excluded fiscal rules).

Table 1.2: **National numerical fiscal rules in the EU countries in our sample**

| Countries | FR (All Fiscal Rules) | BBR (Budget Balance Rules) | DR (Debt Rules) | ER (Expenditure Rules) |
|---|---|---|---|---|
| Bulgaria | 2003-2013 | 2006-2013 | 2003-2013 | 2006-2009 – 2012-2013 |
| Croatia | 2009-2013 | 2012-2013 | 2009-2013 | 2012-2013 |
| Denmark | 2000-2013 | 2000-2013 | - | 2000-2013 |
| Estonia | 2000-2013 | 2000-2013 | - | - |
| Finland | 2000-2013 | 2000-2013 | 2000-2006 – 2010-2013 | 2003-2013 |
| France | 2000-2013 | - | - | 2000-2013 |
| Germany | 2000-2013 | 2000-2013 | - | 2000-2009 – 2012-2013 |
| Greece | 2010-2013 | - | - | 2010-2013 |
| Hungary | 2004-2011 | 2004-2011 | - | 2010-2011 |
| Latvia | 2013 | 2013 | 2013 | - |
| Lithuania | 2000-2013 | - | 2000-2013 | 2008-2013 |
| Luxembourg | 2000-2013 | - | 2000-2013 | 2000-2013 |
| Netherlands | 2000-2013 | - | - | 2000-2013 |
| Poland | 2000-2013 | - | 2000-2013 | 2011-2013 |
| Romania | 2010-2013 | 2013 | 2013 | 2010-2012 |
| Slovak Republic | 2012-2013 | - | 2012-2013 | - |
| Slovenia | 2000-2004 | - | 2000-2004 | - |
| Spain | 2006-2013 | 2006-2013 | - | 2011-2013 |
| Sweden | 2000-2013 | 2000-2013 | - | 2000-2013 |
| United Kingdom | 2000-2008 – 2010-2013 | 2000-2008 – 2010-2013 | 2001-2008 – 2011 -2013 | - |

We measure fiscal rules (FR) through a binary variable that equals one if in a given country for a given year a numerical constraint exists on the national public finance aggregates, namely a budget balance rule (BBR), a debt rule (DR), or an expenditure rule

[9]Since the largest majority of EU countries adopted fiscal rules starting 2013, the treated units become too numerous with respect to the control units for propensity-score matching estimations to be reliable.

(ER). Table 2 summarizes the twenty countries in which there was—at least in a given year during 2000-2013—a fiscal rule corresponding to the definition that we retained for a national numerical rule. For all country-year observations reported in Table 2, the dummy variable capturing the presence of a rule equals 1 if that rule is in place. On the contrary, in country-year observations in which a rule was not in place, the fiscal rule dummy variable equals zero. These observations, together with all the country-year observations for the countries in which there was no rule during the studied period (namely: Austria, Belgium, Cyprus, Czech Republic, Ireland, Italy, Malta, and Portugal), constitute the control group.

## 1.4 Methodology: propensity-score matching and the inverse-probability-weighted regression adjustment

### 1.4.1 The propensity scores matching method

As discussed in the introduction, to estimate the causal effect of fiscal rules on fiscal discipline, we draw upon the propensity score matching method. The goal is to compute the Average Treatment effect on the Treated (ATT), which is defined as the variation in fiscal discipline (Y) in a country that adopted a FR ($Y_1$) had it has not adopted a FR ($Y_0$), namely[10]

$$\text{ATT} = \text{E}[(\text{Y}_1 - \text{Y}_0)|\text{FR} = 1] = \text{E}[\text{Y}_1|\text{FR} = 1] - \text{E}[\text{Y}_0|\text{FR} = 1]. \tag{1.2}$$

Naturally, the problem is that the latter variable $\text{E}[\text{Y}_0|\text{FR} = 1]$ is not observable, and simply comparing the fiscal discipline of the countries that adopted FR with that of countries that did not adopt FR may raise a self-selection issue leading to biased estimates, given that the treatment (i.e. FR adoption) is likely not random. Instead, we compare the fiscal discipline of countries that adopted FR with that of countries that did not adopt FR, but present a close set of observable characteristics X, namely

$$\text{E}[\text{Y}_1|\text{FR} = 1, \text{X}] - \text{E}[\text{Y}_0|\text{FR} = 0, \text{X}]. \tag{1.3}$$

[10] Since matching can be performed only with treatment variables that are binary, it excludes the use of continuous variables such as fiscal rules stringency. In the section devoted to the analysis of heterogeneity we consider various variables that could capture the stringency of fiscal rules (see subsection 7.3 below).

However, as the number of variables in the vector X can be large, Rosenbaum and Rubin [1983] propose to match the treated and untreated units based on their propensity scores, defined by the probability of adoption of the treatment—in our case, the adoption of a fiscal rule—conditional on the vector of observable characteristics X. Assuming that the common support hypothesis ($p(X_i) < 1$, i.e. there exist some comparable control units for each treated unit) is verified—which is the case in our study, as shown by Appendix 2.1—the final expression of the ATT becomes

$$ATT = E[Y_1|FR = 1, p(X)] - E[Y_0|FR = 0, p(X)]. \tag{1.4}$$

### 1.4.2 Computational issues

The computation of the ATT requires knowledge on the propensity scores and the matching method. Regarding the former, we computed the probability of fiscal rules adoption using a vector of characteristics X inspired by existing studies on the determinants of fiscal rules. First, we include the lagged value of CAPB; according to Calderon and Schmidt-Hebbel [2008] and Tapsoba [2012], we expect countries with sound public finance to enact fiscal rules. Second, in the same vein, the lagged value of the debt ratio to real GDP should negatively impact the likelihood of fiscal rules. Third, countries with high real GDP per capita growth rates may benefit of such good conditions to adopt fiscal rules. Fourth, countries with high inflation rates would be less expected to adopt fiscal rules that they may not respect. Fifth, following Guerguil et al. [2017], we include government stability; its effect on fiscal rules is ambiguous, since stable governments could enact fiscal rules to support their policies, but in the same time they may not need such rules given their stability. Sixth, Bonatti and Cristini [2008] showed that the Stability and Growth Pact (SGP) could ensure the coordination of the fiscal policies delegated to the Member States. Therefore, we include a dummy variable to capture the impact of the presence of the SGP on the probability of adopting fiscal rules. Since this SGP dummy variable refers to the European Monetary Union (EMU) accession, which implies that a country is automatically submitted to the SGP, it equals 1 if a country is part of the EMU.[11] Seventh, we control for

[11] Indeed, only eurozone members are under SGP potential sanctions for non-compliance (see Council Regulation (EC) No 1467/97 of 7 July 1997 on speeding up and clarifying the implementation of the excessive deficit procedure); as such, the SGP is a supranational fiscal rule (following Kopits and Symansky [1998] definition) for the eurozone members.

the EU membership through a dummy variable in order to capture the effect of EU accession on the adoption of national fiscal rules.[12] Eighth, we expect a positive relationship between the unemployment rate and the presence of fiscal rules, as a sign of countries' efforts to cope with unemployed population in the EU. Ninth, the appreciation of the real effective exchange rate (REER) may signal good macroeconomic conditions that could support the presence of fiscal rules. Finally, a higher trade openness may signal more open countries that are more exposed to external shocks, and hence more reluctant to adopt fiscal rules that they may not respect.

Regarding the latter, we consider several matching methods for robustness issues. Following Caliendo and Kopeinig [2008], we draw upon five methods, namely: (*i*) the nearest neighbor matching (with N=1 and N=3 neighbors), (*ii*) the radius matching (with a small, a medium, and a large radius, namely: r=0,01, r=0.025 and r=0.05), (*iii*) the kernel matching, (*iv*) the local linear matching, and (*v*) the stratification matching. The N-Nearest Neighbor matches each treated unit with the N control units that have the closest propensity scores; however, if the nearest neighbor is ultimately very far away, this method can suffer from a risk of poor correspondence. By applying a level of tolerance on the maximum distance of the propensity score (a caliper or radius), the Radium Matching method reduces this risk; nevertheless, a limitation of the Radius Matching is that it imposes a brutal cut-off point through the specification of the radius $r$. The non-parametric Kernel Matching deals with this issue, by associating a treated unit with all control units (the counterfactual) weighted proportionately to their proximity to the treated unit. Relatedly, Local Linear Matching is comparable with the Kernel matching, but allows decreasing potential biases related to the estimation of the histogram by drawing upon a linear term in the definition of the weights. Finally, with Stratification Matching, the common support of the propensity score is split into several intervals (or strata); since the effect is computed for each of these strata, this allows accounting for potential heterogeneity in the data.

## 1.5 Benchmark results

### 1.5.1 Fiscal discipline measured by the CAPB

We first present estimations using the traditional measure of fiscal discipline from the related literature, namely the cyclically-adjusted primary balance (CAPB). As illustrated by column (1) of Table 3, the probability of adoption of fiscal rules depends significantly on the past cyclically adjusted primary balances; this finding supports our use of matching to control for reverse causality (see the discussion in Heinemann et al. [2018]). In addition, a higher (lagged) debt ratio, inflation rate, and trade openness are associated with a decrease in the likelihood of fiscal rules adoption, while the opposite holds for government stability, the unemployment rate, and the REER.

Based on propensity scores estimated using column (1) (see Table 3), Table 4 displays the results of the matching. All ATT coefficients reported on line (1) are positive and statistically significant, suggesting that—on average—countries with fiscal rules experience a significant increase of the CAPB with respect to comparable countries that did not adopt fiscal rules. The magnitude of this effect is sizeable, as the improvement of the CAPB (expressed in ratio of GDP) is estimated around 0.5 percentage points (hereafter pp) depending on the considered method of matching.

Moreover, given the debates on the performances of the Hodrick-Prescott filter for the computation of the output gap, we draw upon a trigonometric filter to compute an alternative output gap, and an alternative CAPB measure. Based on propensity scores estimated in column (2) of Table 3, we report on line (2) of Table 4 the ATTs. Despite some significance loss for N=1 nearest neighbor matching, ATTs are positive, significant, and of comparable magnitude with our previous results.

Finally, some authors, e.g Andersen [2013], point out that the residual method may lead to biased estimates of the CAPB, due to the presence of errors and noise in the fiscal variables. Consequently, we perform the matching using propensity scores computed based on the CAPB series calculated by the IMF using the production-function approach (see Girouard and André [2005] and Fedelino et al. [2009]).[13] Based on column (3) of Table

[12]Of course, since not all EU members are equally part of the EMU, the SGP and the EU dummies are not collinear.

[13]We report that the use of IMF's CAPB measure is equally motivated by the lack of availability of the European Commission's measure of the CAPB for the beginning of our period of study.

3, ATTs reported on line (3) of Table 4 are—although of a higher magnitude—yet again consistent with our previous findings.

Overall, our results contribute to the debate regarding the effect of fiscal rules on fiscal discipline measured by the CAPB, by revealing—based on a treatment effect analysis that tackles potential endogeneity in the presence of fiscal rules—a favorable effect in our sample of EU countries. In particular, the magnitude of this effect is somewhat weaker for the EU countries with respect to the developing countries (see the results in Tapsoba [2012], who employs the same methodology).

Table 1.4: **Matching Results: ATT of FR on the CAPB**

| | Nearest-neighbor Matching | | Stratification Matching | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Dependent variable: $CAPB_{i,t}$ calculated using the Hodrick-Prescott (HP) filter | | | | | | | | | |
| **[1] ATT** | **0.698**$^{**}$ | **0.451**$^{*}$ | **0.549**$^{***}$ | **0.676**$^{***}$ | **0.540**$^{**}$ | **0.515**$^{**}$ | **0.517**$^{***}$ | **0.510**$^{**}$ | **0.290**$^{**}$ |
| | (0.340) | (0.212) | (0.203) | (0.273) | (0.222) | (0.212) | (0.206) | (0.207) | (0.150) |
| Number of treated observations | 203 | 203 | 203 | 191 | 203 | 203 | 203 | 203 | 203 |
| Number of control observations | 188 | 188 | 188 | 188 | 188 | 188 | 180 | 188 | 188 |
| Dependent variable: $CAPB_{i,t}$ calculated using the trigonometric filter | | | | | | | | | |
| **[2] ATT** | **0.734**$^{**}$ | **0.777**$^{***}$ | **0.544**$^{**}$ | **0.712**$^{**}$ | **0.518**$^{**}$ | **0.546**$^{**}$ | **0.542**$^{***}$ | **0.537**$^{***}$ | **0.288**$^{**}$ |
| | (0.308) | (0.306) | (0.204) | (0.328) | (0.210) | (0.261) | (0.207) | (0.211) | (0.150) |
| Dependent variable: $CAPB_{i,t}$ calculated using the production function approach, source IMF | | | | | | | | | |
| **[3] ATT** | **1.341**$^{***}$ | **1.459**$^{***}$ | **0.640**$^{***}$ | **1.243**$^{***}$ | **1.424**$^{***}$ | **1.389**$^{***}$ | **1.481**$^{***}$ | **1.378**$^{***}$ | **0.942**$^{***}$ |
| | (0.478) | (0.501) | (0.205) | (0.383) | (0.363) | (0.317) | (0.379) | (0.365) | (0.282) |

Note: Bootstrapped standard errors (with 500 replications) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. For stratification matching the number of strata is five and the level of significance is 0.01. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment. IPWRA includes all control variables for propensity scores estimation step.

### 1.5.2 A novel measure of fiscal discipline: the Global Fiscal Performance Index (GFPI)

We now look at the effect of fiscal rules on our new measure of fiscal discipline—the global fiscal performance index. Based on propensity scores estimated using model (1) in Table 3, the line (1) of Table 5 reports the ATTs. Results are comparable with those based on the CAPB: countries with fiscal rules present higher values of the GFPI with respect to comparable countries that did not adopt fiscal rules. In particular, the presence of a fiscal rule improves the GFPI on average by around 0.5 units, an economically-meaningful effect given the range of the GFPI values in our sample (between -2 and 4).

Aside from the comparable effect of fiscal rules on the CAPB and the GFPI, it would be interesting to observe their effect on the components of the GFPI. As illustrated by lines (2)-(6) of Table 5, the effect of fiscal rules on the different GFPI components is fairly different. First, irrespective of the matching method, fiscal rules are found to significantly reduce both public and external deficits—see lines (2) and (6). Second, the favorable effect of fiscal rules on the debt growth rate is significant for all but one matching methods, and for all but three matching methods when considering the interest growth rate, as shown by lines (3) and (4). Finally, fiscal rules are not found to exert a significant effect on the Variations of Fiscal Revenues (considered with a negative sign), irrespective of the matching method. These results show that the impact of fiscal rules on fiscal discipline varies depending on the way fiscal discipline is measured, and—therefore—justifies our strategy of capturing fiscal discipline in several ways. The next section analyzes the robustness of our findings.

Table 1.5: **Matching Results: ATT of FR on the GFPI, and its components**

| | Nearest-neighbor Matching | | Stratification Matching | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Dependent variable: GFPI$_{i,t}$ | | | | | | | | | |
| **[1] ATT** | **0.494**$^{***}$ | **0.537**$^{***}$ | **0.550**$^{***}$ | **0.525**$^{***}$ | **0.504**$^{***}$ | **0.490**$^{***}$ | **0.499**$^{***}$ | **0.491**$^{***}$ | **0.351**$^{***}$ |
| | (0.162) | (0.149) | (0.203) | (0.142) | (0.127) | (0.110) | (0.119) | (0.104) | (0.105) |
| Number of treated observations | 203 | 203 | 203 | 191 | 203 | 203 | 203 | 203 | 203 |
| Number of control observations | 188 | 188 | 188 | 188 | 188 | 188 | 180 | 188 | 188 |
| Dependent variable: $Public\ Deficit_{i,t}$ | | | | | | | | | |
| **[2] ATT** | **-1.953**$^{***}$ | **-1.822**$^{***}$ | **-1.772**$^{***}$ | **-2.111**$^{***}$ | **-2.036**$^{***}$ | **-1.907**$^{***}$ | **-2.044**$^{***}$ | **-1.894**$^{***}$ | **-1.942**$^{***}$ |
| | (0.628) | (0.478) | (0.312) | (0.471) | (0.423) | (0.398) | (0.354) | (0.399) | (0.267) |
| Dependent variable: $Debt\ growth\ rate_{i,t}$ | | | | | | | | | |
| **[3] ATT** | **-3.007**$^{***}$ | **-3.602**$^{*}$ | **-4.550**$^{***}$ | **-4.227**$^{*}$ | **-2.795** | **-4.319**$^{**}$ | **-5.828**$^{***}$ | **-4.349**$^{*}$ | **-5.600**$^{***}$ |
| | (3.529) | (2.387) | (1.273) | (2.120) | (2.443) | (1.952) | (1.716) | (1.940) | (1.472) |
| Dependent variable: $Interest\ growth\ rate_{i,t}$ | | | | | | | | | |
| **[4] ATT** | **-8.603**$^{*}$ | **-6.175** | **-4.980**$^{***}$ | **-8.209**$^{**}$ | **-7.198**$^{*}$ | **-5.320** | **-6.488**$^{*}$ | **-5.817** | **-6.170**$^{**}$ |
| | (5.608) | (4.692) | (1.825) | (3.973) | (4.429) | (4.747) | (3.649) | (4.204) | (2.722) |
| Dependent variable: $Variations\ of\ fiscal\ revenues_{i,t}$ | | | | | | | | | |
| **[5] ATT** | **-0.006** | **-0.005** | **-0.183** | **-0.042** | **0.051** | **-0.107** | **-0.060** | **-0.054** | **0.149** |
| | (0.625) | (0.570) | (0.324) | (0.484) | (0.510) | (0.467) | (0.452) | (0.497) | (0.401) |
| Dependent variable: $External\ Deficit_{i,t}$ | | | | | | | | | |
| **[6] ATT** | **-22.742**$^{***}$ | **-16.831**$^{***}$ | **-12.000**$^{***}$ | **-19.105**$^{***}$ | **-18.043**$^{***}$ | **-16.495**$^{***}$ | **-17.077**$^{***}$ | **-16.816**$^{***}$ | **-14.215**$^{***}$ |
| | (6.608) | (5.039) | (4.141) | (5.151) | (5.672) | (3.763) | (4.908) | (4.634) | (3.752) |

Note: Bootstrapped standard errors (with 500 replications) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. For stratification matching the number of strata is five and the level of significance is 0.01. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment. IPWRA includes all control variables for propensity scores estimation step except for ATT [5] and [6] where we removed the dependent variable from control variables included in propensity scores estimation.

Table 1.3: **Probit estimates of the Propensity Scores**

| Dependent variable: FR | [1] | [2] | [3] |
|---|---|---|---|
| CAPB computed with the: | HP Filter | Trigonometric Filter | IMF Production Function |
| Intercept | -0.983 | -0.928 | -3.341** |
| | (1.048) | (1.171) | (1.319) |
| $CAPB_{t-1}$ | 0.113*** | 0.107*** | 0.127*** |
| | (0.042) | (0.042) | (0.029) |
| Debt $ratio_{t-1}$ | -0.018*** | -0.018*** | -0.019*** |
| | (0.003) | (0.003) | (0.003) |
| Real per capita GDP growth rate | -0.028 | -0.032 | -0.042* |
| | (0.021) | (0.021) | (0.023) |
| Inflation rate | -0.103*** | -0.102*** | -0.045* |
| | (0.026) | (0.026) | (0.023) |
| Government stability | 0.065* | 0.067* | 0.243* |
| | (0.200) | (0.200) | (0.222) |
| SGP | -0.080 | -0.083 | -0.161 |
| | (0.159) | (0.162) | (0.171) |
| Dummy EU membership | 0.077 | 0.070 | 0.015 |
| | (0.386) | (0.386) | (0.455) |
| Unemployment rate | 0.030* | 0.029* | 0.075*** |
| | (0.019) | (0.019) | (0.022) |
| REER | 0.026*** | 0.026*** | 0.047*** |
| | (0.009) | (0.009) | (0.011) |
| Trade openness | -0.008** | -0.008** | -0.010** |
| | (0.003) | (0.003) | (0.003) |
| Adjusted $R^2$ | 0.138 | 0.137 | 0.175 |
| Observations | 392 | 392 | 392 |

Note: Robust standard errors in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively.

## 1.6 Robustness

We investigate the robustness of our results in several ways.

First, (*i*) following Rosenbaum and Rubin [1985], we analyze the conditional independence assumption, i.e. the absence of significant differences between the observable characteristics of the treated and non-treated observations. To this end, we look at the absolute standardized mean difference between observations with and without fiscal rules. The results of the equality test of the mean difference (standardized bias) between the observables of FRers and non-FRers returns high p-values, namely above 0.1 in all cases when using the CAPB (see Table 6, below the line (1) that recalls the benchmark estimations) or the GFPI (see Table 7, below the line (1) that recalls the benchmark estimations). Consequently, there are no statistical differences between the two groups after matching, which supports the efficiency of our matching procedure. (*ii*) Moreover, following e.g. Guerguil et al. [2017], we use the Rosenbaum [2002] bounding sensitivity test to check whether unobserved heterogeneity affects our results;[14] Appendix 12 shows that our results are not biased by unobserved factors. (*iii*) Lastly, in our main estimations we use plain bootstrapping to compute standard errors; although appealing for our relatively small sample, this may raise an overfitting issue with detrimental consequences for the generality of the results. However, as shown by Appendix 8 and Appendix 9, the use of cluster-bootstrapped standard errors (see the discussion in Bertrand et al. [2004]) leaves our results unaffected.

Second, to see if our findings are specific to the propensity-score matching method, we draw upon the inverse-probability-weighted regression adjustment (IPWRA) estimator, which uses coefficients from a weighted regression to obtain averages of treatment-level predicted outcomes. The weights come from the estimated inverse probabilities of treatment, and the treatment effects correspond to the contrasts of the averages. This estimator is considered as a doubly robust estimator: it is robust to a potential misspecification bias in the propensity score, and is not sensitive to the sample size (see e.g. Imbens and Wooldridge [2009] for a comprehensive review of the method). IPWRA estimations—with all the matching variables in the outcome equation— reported in the last columns of Table 4 (for the CAPB) and Table 5 (for the GFPI) confirm that, except for some magnitude

[14] An important source of unobserved heterogeneity is related to voters' preferences for fiscal discipline, see e.g. Poterba [1996] and Krogstrup and Walti [2008].

loss, fiscal rules improve fiscal discipline irrespective of the way it is being measured.

Third, some countries may be involved into the use of fiscal gimmickry—see e.g. Alt et al. [2014] for a discussion. Such a behaviour may potentially affect the effect of rules on discipline—and even whether a rule is adopted in the first place. To explore this issue, we follow the strategy retained by Alt et al. [2014], and draw upon a measure of stock-flow adjustments of public debt. Using propensity scores computed based on the probit model in column (2) of Appendix 4, the line (2) in Table 6 and Table 7 show that all ATTs are significant and of a comparable magnitude with our benchmark findings.

Fourth, we investigate the impact of additional control variables in two ways. (*i*) Following Tapsoba [2012], we consider an additional set of control variables in the probit specification, namely: external deficit, the Variations of Fiscal Revenues, the output gap, the lagged squared debt ratio, the government fragmentation, a dummy variable for the presence of elections, a dummy variable for emerging countries, and a dummy variable indicating if there was a reform of the SGP—2005, 2011 (the Six Pack), and 2013 (the Two Pack). Based on propensity scores computed using the probit models from columns (3)-(10) in Appendix 4, Table 6 and Table 7 report the ATTs for the CAPB and GFPI, respectively, and confirm the robustness of our benchmark results, both in significance and magnitude. (*ii*) In addition, we estimate a saturated propensity score model that includes all the additional matching variables at the same time (see e.g. Jorda and Taylor [2016]). As shown by the line (11) of Table 6 and Table 7, the presence of fiscal rules significantly increases both CAPB and the GFPI in all specifications; although we observe some magnitude loss in the effect of fiscal rules on fiscal discipline, these results provide additional support for our modeling strategy with respect to the unobservables.

Finally, we perform estimations on the sub-sample of core EU countries, by excluding the new EU countries, i.e. that entered the EU after 2004, and Greece. Using propensity scores computed using the last column of Appendix 4, we reveal in the last line of Table 6 (for CAPB) and Table 7 (for GFPI) ATTs that support—yet again—a favorable effect of fiscal rules on fiscal discipline. Nevertheless, compared with our previous findings, the significance of the effect is weaker for the CAPB (only in six out of eight cases), and its magnitude stronger for both CAPB and GFPI. Such differences motivate the next section, devoted to the analysis of possible heterogeneities in the effect of fiscal rules on fiscal discipline.

Table 1.6: **Matching Results: ATT of FR on the CAPB—Robustness**

| | Nearest-neighbor Matching | | Stratification Matching | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Dependent variable: $CAPB_{i,t}$ calculated with the Hodrick-Prescott (HP) filter | | | | | | | | | |
| **[1] ATT-CAPB** | **0.698**$^{*}$ | **0.451**$^{*}$ | **0.549**$^{***}$ | **0.676**$^{***}$ | **0.540**$^{**}$ | **0.515**$^{**}$ | **0.517**$^{***}$ | **0.510**$^{**}$ | **0.290**$^{**}$ |
| | (0.340) | (0.212) | (0.203) | (0.273) | (0.222) | (0.212) | (0.206) | (0.207) | (0.150) |
| Number of treated observations | 203 | 203 | 203 | 191 | 203 | 203 | 203 | 203 | 203 |
| Number of control observations | 188 | 188 | 188 | 188 | 188 | 188 | 180 | 188 | 188 |
| Standardized bias (p-value) | 0.628 | 0.898 | 0.262 | 0.714 | 0.992 | 0.997 | 0.628 | 0.997 | - |
| **[2] SFA on debt** | **0.580**$^{**}$ | **0.608**$^{*}$ | **0.509**$^{***}$ | **0.571**$^{*}$ | **0.584**$^{***}$ | **0.497**$^{**}$ | **0.489**$^{*}$ | **0.536**$^{**}$ | **0.367**$^{**}$ |
| | (0.314) | (0.244) | (0.201) | (0.342) | (0.225) | (0.233) | (0.297) | (0.235) | (0.166) |
| **[3] Adding external deficit** | **0.518**$^{*}$ | **0.468**$^{*}$ | **0.554**$^{***}$ | **0.499**$^{**}$ | **0.526**$^{**}$ | **0.487**$^{**}$ | **0.444**$^{**}$ | **0.487**$^{**}$ | **0.391**$^{***}$ |
| | (0.280) | (0.275) | (0.210) | (0.242) | (0.252) | (0.200) | (0.253) | (0.195) | (0.171) |
| **[4] Adding Variations of Fiscal Revenues** | **0.521**$^{*}$ | **0.413**$^{*}$ | **0.504**$^{***}$ | **0.393**$^{*}$ | **0.393**$^{*}$ | **0.453**$^{**}$ | **0.472**$^{**}$ | **0.455**$^{**}$ | **0.386**$^{**}$ |
| | (0.327) | (0.315) | (0.193) | (0.305) | (0.247) | (0.239) | (0.215) | (0.237) | (0.169) |
| **[5] Adding output gap** | **0.332** | **0.609**$^{*}$ | **0.554**$^{***}$ | **0.555**$^{*}$ | **0.531**$^{*}$ | **0.535**$^{**}$ | **0.522**$^{***}$ | **0.515**$^{**}$ | **0.415**$^{***}$ |
| | (0.307) | (0.329) | (0.203) | (0.383) | (0.249) | (0.198) | (0.200) | (0.263) | (0.175) |
| **[6] Adding lagged squared debt** | **0.703**$^{**}$ | **0.576**$^{**}$ | **0.524**$^{***}$ | **0.791**$^{***}$ | **0.556**$^{***}$ | **0.485**$^{***}$ | **0.522**$^{**}$ | **0.507**$^{**}$ | **0.385**$^{**}$ |
| | (0.319) | (0.285) | (0.200) | (0.299) | (0.215) | (0.242) | (0.232) | (0.258) | (0.174) |
| **[7] Adding gov. fragmentation** | **0.286** | **0.347**$^{*}$ | **0.522**$^{***}$ | **0.455**$^{*}$ | **0.493**$^{**}$ | **0.510**$^{**}$ | **0.516**$^{**}$ | **0.511**$^{**}$ | **0.395**$^{**}$ |
| | (0.389) | (0.254) | (0.205) | (0.334) | (0.280) | (0.227) | (0.254) | (0.255) | (0.173) |
| **[8] Adding electoral system** | **0.825**$^{**}$ | **0.616**$^{*}$ | **0.533**$^{***}$ | **0.706**$^{*}$ | **0.577**$^{***}$ | **0.491**$^{**}$ | **0.473**$^{**}$ | **0.505**$^{**}$ | **0.389**$^{**}$ |
| | (0.403) | (0.295) | (0.202) | (0.412) | (0.231) | (0.236) | (0.248) | (0.258) | (0.176) |
| **[9] Adding emerging country** | **0.416** | **0.468** | **0.550**$^{***}$ | **0.663**$^{**}$ | **0.536**$^{**}$ | **0.507**$^{**}$ | **0.516**$^{**}$ | **0.512**$^{**}$ | **0.415**$^{***}$ |
| | (0.292) | (0.337) | (0.204) | (0.304) | (0.288) | (0.221) | (0.226) | (0.263) | (0.175) |
| **[10] Adding PSC reforms** | **0.346** | **0.626**$^{**}$ | **0.549**$^{***}$ | **0.551**$^{**}$ | **0.534**$^{**}$ | **0.518**$^{**}$ | **0.489**$^{**}$ | **0.518**$^{**}$ | **0.425**$^{**}$ |
| | (0.323) | (0.317) | (0.206) | (0.246) | (0.306) | (0.241) | (0.229) | (0.260) | (0.182) |
| **[11] Satured PS** | **0.478**$^{*}$ | **0.377**$^{*}$ | **0.442**$^{**}$ | **0.405**$^{*}$ | **0.349**$^{*}$ | **0.311**$^{*}$ | **0.290**$^{**}$ | **0.311**$^{*}$ | **0.290**$^{**}$ |
| | (0.263) | (0.238) | (0.203) | (0.228) | (0.273) | (0.266) | (0.213) | (0.197) | (0.150) |
| **[12] Excl. New EU & Greece** | **1.208**$^{*}$ | **1.268** | **1.149**$^{***}$ | **1.250** | **1.147**$^{*}$ | **1.214**$^{*}$ | **1.346**$^{*}$ | **1.216**$^{*}$ | **1.406**$^{***}$ |
| | (0.893) | (0.863) | (1.203) | (0.961) | (1.147) | (0.744) | (1.346) | (0.747) | (0.425) |

Note: Bootstrapped standard errors (with 500 replications) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. For stratification matching the number of strata is five and the level of significance is 0.01. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment. IPWRA includes all control variables for propensity scores estimation step. SFA on debt stands for Stock Flow Adjustement on general government gross debt. Saturated PS stands for Saturated Propensity Scores which includes all control variables in the propensity score.

Table 1.7: **Matching Results: ATT of FR on the GFPI—Robustness**

| | Nearest-neighbor Matching | | Stratification Matching | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Dependent variable: $\text{GFPI}_{i,t}$ | | | | | | | | | |
| **[1] ATT-GFPI** | **0.494**$^{***}$ | **0.537**$^{***}$ | **0.550**$^{***}$ | **0.525**$^{***}$ | **0.504**$^{***}$ | **0.490**$^{***}$ | **0.499**$^{***}$ | **0.491**$^{***}$ | **0.351**$^{***}$ |
| | (0.162) | (0.149) | (0.203) | (0.142) | (0.127) | (0.110) | (0.119) | (0.104) | (0.105) |
| Number of treated observations | 203 | 203 | 203 | 191 | 203 | 203 | 203 | 203 | 203 |
| Number of control observations | 188 | 188 | 188 | 188 | 188 | 188 | 180 | 188 | 188 |
| Standardized bias (p-value) | 0.628 | 0.898 | 0.262 | 0.714 | 0.992 | 0.997 | 0.628 | 0.997 | - |
| **[2] SFA on debt** | **0.494**$^{***}$ | **0.540**$^{***}$ | **0.468**$^{***}$ | **0.509**$^{***}$ | **0.498**$^{***}$ | **0.521**$^{***}$ | **0.514**$^{***}$ | **0.519**$^{***}$ | **0.441**$^{***}$ |
| | (0.159) | (0.167) | (0.104) | (0.136) | (0.137) | (0.136) | (0.121) | (0.144) | (0.118) |
| **[3] Adding external deficit** | **0.396**$^{**}$ | **0.348**$^{**}$ | **0.366**$^{***}$ | **0.457**$^{***}$ | **0.433**$^{***}$ | **0.389**$^{**}$ | **0.392**$^{***}$ | **0.391**$^{**}$ | **0.374**$^{***}$ |
| | (0.209) | (0.173) | (0.097) | (0.166) | (0.124) | (0.181) | (0.123) | (0.157) | (0.125) |
| **[4] Adding Variations of Fiscal Revenues** | **0.393**$^{**}$ | **0.435**$^{***}$ | **0.446**$^{***}$ | **0.452**$^{***}$ | **0.466**$^{***}$ | **0.475**$^{***}$ | **0.463**$^{***}$ | **0.471**$^{***}$ | **0.398**$^{***}$ |
| | (0.184) | (0.122) | (0.094) | (0.124) | (0.146) | (0.127) | (0.103) | (0.140) | (0.112) |
| **[5] Adding output gap** | **0.437**$^{**}$ | **0.532**$^{***}$ | **0.475**$^{***}$ | **0.532**$^{***}$ | **0.589**$^{***}$ | **0.503**$^{***}$ | **0.500**$^{***}$ | **0.496**$^{***}$ | **0.435**$^{***}$ |
| | (0.175) | (0.143) | (0.102) | (0.166) | (0.136) | (0.139) | (0.122) | (0.113) | (0.122) |
| **[6] Adding lagged squared debt** | **0.537**$^{***}$ | **0.554**$^{***}$ | **0.461**$^{***}$ | **0.546**$^{***}$ | **0.510**$^{***}$ | **0.493**$^{***}$ | **0.505**$^{***}$ | **0.497**$^{***}$ | **0.424**$^{***}$ |
| | (0.172) | (0.159) | (0.102) | (0.129) | (0.168) | (0.163) | (0.122) | (0.125) | (0.126) |
| **[7] Adding gov. fragmentation** | **0.534**$^{***}$ | **0.492**$^{***}$ | **0.459**$^{***}$ | **0.501**$^{***}$ | **0.520**$^{***}$ | **0.491**$^{***}$ | **0.521**$^{***}$ | **0.494**$^{***}$ | **0.407**$^{***}$ |
| | (0.173) | (0.147) | (0.102) | (0.133) | (0.115) | (0.123) | (0.121) | (0.115) | (0.129) |
| **[8] Adding electoral system** | **0.445**$^{***}$ | **0.606**$^{***}$ | **0.489**$^{***}$ | **0.579**$^{***}$ | **0.538**$^{***}$ | **0.497**$^{***}$ | **0.490**$^{***}$ | **0.501**$^{***}$ | **0.419**$^{***}$ |
| | (0.171) | (0.138) | (0.103) | (0.152) | (0.142) | (0.144) | (0.126) | (0.167) | (0.127) |
| **[9] Adding emerging country** | **0.535**$^{***}$ | **0.533**$^{***}$ | **0.470**$^{***}$ | **0.511**$^{***}$ | **0.493**$^{***}$ | **0.487**$^{***}$ | **0.499**$^{***}$ | **0.488**$^{***}$ | **0.433**$^{***}$ |
| | (0.190) | (0.153) | (0.103) | (0.155) | (0.152) | (0.123) | (0.120) | (0.109) | (0.123) |
| **[10] Adding PSC reforms** | **0.544**$^{***}$ | **0.521**$^{***}$ | **0.476**$^{***}$ | **0.486**$^{***}$ | **0.502**$^{***}$ | **0.501**$^{***}$ | **0.492**$^{***}$ | **0.500**$^{***}$ | **0.429**$^{***}$ |
| | (0.150) | (0.144) | (0.103) | (0.137) | (0.139) | (0.135) | (0.127) | (0.118) | (0.120) |
| **[11] Saturated PS** | **0.343**$^{**}$ | **0.386**$^{***}$ | **0.341**$^{***}$ | **0.367**$^{***}$ | **0.378**$^{***}$ | **0.368**$^{***}$ | **0.382**$^{***}$ | **0.296**$^{***}$ | **0.351**$^{***}$ |
| | (0.174) | (0.117) | (0.085) | (0.124) | (0.144) | (0.132) | (0.098) | (0.373) | (0.105) |
| **[12] Excl. New EU & Greece** | **0.759**$^{***}$ | **0.703**$^{***}$ | **0.670**$^{***}$ | **0.678**$^{**}$ | **0.682**$^{***}$ | **0.717**$^{***}$ | **0.729**$^{***}$ | **0.713**$^{***}$ | **0.650**$^{***}$ |
| | (0.201) | (0.247) | (0.200) | (0.296) | (0.228) | (0.242) | (0.202) | (0.209) | (0.158) |

Note: Bootstrapped standard errors (with 500 replications) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. For Stratification matching, the number of strata is five and the level of significance is 0.01. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment. IPWRA includes all control variables for propensity scores estimation step except for ATT [2] and [3] where we removed the dependent variable from control variables included in propensity scores estimation. SFA on debt stands for Stock Flow Adjustement on general government gross debt. Saturated PS stands for Saturated Propensity Scores, the propensity score specification includes all control variables.

## 1.7 Heterogeneity in fiscal rules performance

### 1.7.1 The type of fiscal rule

So far, our analysis focused on the effect of fiscal rules altogether. In the following, based on the propensity scores estimated using the columns (1)-(3) in Appendix 5, we look at the effect of the different types of fiscal rules, namely, budget balance rules (BBR) in Table 8, expenditure rules (ER) in Table 9, and debt rules (DR) in Table 10,[15] on fiscal discipline. Prior to discussing the results in detail, we report that the common support hypothesis is verified for each type of fiscal rule (see Appendix 2.2, Appendix 2.3, and Appendix 2.4); the high p-values of the standardized bias test support the conditional independence assumption (see Tables 8, 9, and 10); and using the inverse-probability-weighted regression adjustment estimator confirms our findings based on propensity scores matching (see the last columns of Tables 8, 9, and 10).

Regarding the traditional measure of fiscal discipline, the line (1) in Tables 8, 9, and 10 presents the effects of the different types of fiscal rules on the CAPB. We reveal two important effects. First, the presence of BBR significantly improves the CAPB with respect to comparable countries without BBR. The magnitude of this effect is economically meaningful, around 0.4-0.5 pp, and comparable with our findings when considering all fiscal rules together. Second, neither ER nor DR make a significant difference in terms of fiscal discipline when measured by the CAPB. While the lack of effect of ER on the CAPB is consistent with the conclusions of previous studies, including e.g. Debrun et al. [2008] or Bergman et al. [2016], the absence of a significant effect of DR on the CAPB is more novel with respect to existing studies; for example, DR are associated with a significantly higher CAPB when combined with BBR in Debrun et al. [2008], or by themselves in Bergman et al. [2016]; nevertheless, while Bergman et al. [2016] look at the effect of a strengthening DR that were already in place, our estimates refer to the presence of fiscal rules with respect to their absence. A possible explanation of our findings is that all EU countries are already subject to the 60% debt rule of the SGP—that they mostly fail to respect during the period that we study, contrary to the 3% deficit budget rule that is more closely followed by EU monitoring agencies and therefore more respected—so they have little incentives to respect their national DR.

[15]Due to the low number of countries that enacted Revenue Rules (Denmark, Lithuania, and the Netherlands), we decided not to present the results of their effect on fiscal discipline.

Let us now look at our novel measure of fiscal discipline, namely the GFPI. ATTs reported on the line (2) of Tables 8, 9, and 10 reveal a fairly different picture in the effects of fiscal rules on GFPI compared with the CAPB, on two grounds. First, in addition to BBR, ER significantly improve the GFPI; therefore, the effect of ER on fiscal discipline crucially depends on the way it is being measured, since the presence of ER can either make no statistical difference (when measured by the CAPB) or significantly improve it (when measured by the GFPI). Second, while the size of the effect of BBR on the CAPB was comparable to the size of the effect of fiscal rules altogether, differences in magnitude are at work when considering the GFPI index; indeed, compared with the effect of fiscal rules altogether, estimated around 0.5 units, the impact of BBR on the GFPI is higher (the estimated ATTs are around 0.7 units), and this is also the case for the effect of ER (the estimated ATTs are around 0.6 units).

Given such differences in the effect of fiscal rules on CAPB and GFPI, we examine their impact on the variables composing the GFPI. First, as shown by the line (3) of Tables 8, 9, and 10, similar to the effect of fiscal rules altogether, the presence of BBR or ER yields significantly lower public deficits (with no significant effect of DR). The magnitude of this favorable effect is slightly higher on average for ER (around 2.1 pp) compared with fiscal rules altogether or BBR (around 1.8-2.1 pp). Second, the significance of the effect of BBR and ER on the growth of public debt is comparable with that of fiscal rules altogether—seven (six) out of eight ATTs are significant for BBR (ER), with no significant impact of DR (see the line (4) of Tables 8, 9, and 10). Similar to public deficits, the growth of public debt responds slightly more to ER (around 4-5 pp) and much more to BBR (around 5-6 pp) compared with its response to fiscal rules altogether (around 4 pp). Third, contrary to their significant effect on public deficit and the growth of public debt, ER—similar to DR—do not significantly affect the growth of interest rate (see the line (5) of Tables 9 and 10). However, while the effect of fiscal rules altogether was not found to be significant, the presence of BBR significantly decreases the growth of interest rates (all eight ATTs are significant) by around 7 pp (see the line (5) of Table 8). Fourth, similar to the lack of a significant effect of fiscal rules altogether, the estimated ATTs of the impact of BBR, ER, and DR on the Variations of Fiscal Revenues are not statistically significant, as illustrated by the line (6) of Tables 8, 9, and 10. Fifth, compared with ATTs estimated roughly between 12 and 23 pp for fiscal rules altogether, all ATTs of the effect of BBR on the external deficit are not significant (see the line (7) in Table 8). Instead, half of ATTs are significant for DR (see the line (7) in Table 10), with an estimated effect around 10-14 pp, while the impact of ER is particularly robust—all eight ATTs are significant on the line (7) in Table 9—and around 75% higher compared with the impact of fiscal rules altogether (around 28-30 pp).

Overall, our results show that—contrary to the lack of significant impact of DR—the effect of BBR and ER differs both in significance and magnitude compared with the impact of fiscal rules altogether, depending on the considered fiscal rule and fiscal variable (except for the Variations of Fiscal Revenues, which was not found to be significantly affected). The latter finding may be explained by the fact that none of the various types of fiscal rules explicitly targets fiscal revenues, while the former suggests that debt rules are not sufficiently binding to trigger an improvement in fiscal discipline (in particular, they are found not to significantly impact even the growth of public debt). In addition, the differentiated impact of BBR and ER may be the consequence of the various fiscal aggregates targeted by the two rules, namely the fiscal balance and government expenditure; in particular, aside from differences in the magnitude of their effect on, e.g. deficit or debt, only BBR—that place a direct constraint on the fiscal balance—are judged to be significantly binding to affect the growth of interest rates, probably through changes in expectations about public debt sustainability (see e.g. Badinger and Reuter [2017]).

Table 1.8: **Matching Results with BBR (Budget Balance Rules) as the treatment variable**

| | Nearest-neighbor Matching | | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Dependent variable: $\text{CAPB}_{i,t}$ | | | | | | | | |
| **[1] ATT** | **0.297**$^{*}$ | **0.465**$^{**}$ | **0.423**$^{**}$ | **0.546**$^{***}$ | **0.524**$^{***}$ | **0.503**$^{***}$ | **0.532**$^{***}$ | **0.344**$^{**}$ |
| | (0.273) | (0.232) | (0.207) | (0.211) | (0.195) | (0.182) | (0.205) | (0.166) |
| Number of treated observations | 108 | 108 | 104 | 108 | 108 | 108 | 108 | 116 |
| Number of control observations | 276 | 276 | 276 | 276 | 276 | 276 | 276 | 276 |
| Standardized bias (p-value) | 0.919 | 0.796 | 0.974 | 0.935 | 0.961 | 0.919 | 0.958 | - |
| Dependent variable: $\text{GFPI}_{i,t}$ | | | | | | | | |
| **[2] ATT** | **0.679**$^{***}$ | **0.715**$^{***}$ | **0.717**$^{***}$ | **0.703**$^{***}$ | **0.698**$^{***}$ | **0.676**$^{***}$ | **0.698**$^{***}$ | **0.644**$^{***}$ |
| | (0.184) | (0.151) | (0.153) | (0.131) | (0.134) | (0.135) | (0.137) | (0.136) |
| Dependent variable: $\text{P}ublic\ \text{D}eficit_{i,t}$ | | | | | | | | |
| **[3] ATT** | **-1.953**$^{***}$ | **-1.822**$^{***}$ | **-2.111**$^{***}$ | **-2.036**$^{***}$ | **-1.908**$^{***}$ | **-2.045**$^{***}$ | **-1.942**$^{***}$ | **-2.456**$^{***}$ |
| | (0.549) | (0.402) | (0.424) | (0.374) | (0.475) | (0.386) | (0.385) | (0.287) |
| Dependent variable: $\text{D}ebt\ growth\ rate_{i,t}$ | | | | | | | | |
| **[4] ATT** | **-4.832** | **-5.429**$^{**}$ | **-5.188**$^{**}$ | **-6.180**$^{**}$ | **-6.176**$^{**}$ | **-6.123**$^{***}$ | **-5.855**$^{**}$ | **-4.660**$^{***}$ |
| | (3.583) | (2.501) | (2.669) | (3.248) | (2.713) | (2.372) | (3.164) | (1.647) |
| Dependent variable: $\text{I}nterest\ growth\ rate_{i,t}$ | | | | | | | | |
| **[5] ATT** | **-8.380**$^{*}$ | **-7.676**$^{**}$ | **-7.263**$^{**}$ | **-8.104**$^{**}$ | **-8.029**$^{**}$ | **-7.527**$^{**}$ | **-7.534**$^{**}$ | **-5.436**$^{**}$ |
| | (4.717) | (3.548) | (3.922) | (3.675) | (3.601) | (3.628) | (3.297) | (2.448) |
| Dependent variable: $\text{V}ariations\ of\ fiscal\ revenues_{i,t}$ | | | | | | | | |
| **[6] ATT** | **-0.006** | **-0.005** | **-0.042** | **0.051** | **-0.107** | **-0.060** | **-0.054** | **0.014** |
| | (0.600) | (0.533) | (0.498) | (0.219) | (0.540) | (0.385) | (0.424) | (0.418) |
| Dependent variable: $\text{E}xternal\ \text{D}eficit_{i,t}$ | | | | | | | | |
| **[7] ATT** | **-3.228** | **-3.284** | **-5.077** | **-3.825** | **-3.869** | **-3.798** | **-3.818** | **-6.806** |
| | (7.672) | (6.280) | (6.060) | (5.224) | (5.507) | (5.170) | (4.915) | (4.184) |

Note: Bootstrapped standard errors (with 500 replications) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment. IPWRA includes all control variables for propensity scores estimation step except for ATT [6] and [7] where we removed the dependent variable from control variables included in propensity scores estimation.

Table 1.9: **Matching Results with ER (Expenditure Rules) as the treatment variable**

| | Nearest-neighbor Matching | | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Dependent variable: $CAPB_{i,t}$ | | | | | | | | |
| **[1] ATT** | **0.571** | **0.267** | **0.264** | **0.285** | **0.394** | **0.371** | **0.386** | **0.151** |
| | (0.397) | (0.384) | (0.298) | (0.282) | (0.286) | (0.289) | (0.303) | (0.163) |
| Number of treated observations | 121 | 121 | 117 | 120 | 121 | 121 | 121 | 122 |
| Number of control observations | 270 | 270 | 270 | 270 | 270 | 270 | 270 | 270 |
| Standardized bias (p-value) | 0.898 | 0.977 | 0.953 | 0.954 | 0.989 | 0.898 | 0.987 | - |
| Dependent variable: $GFPI_{i,t}$ | | | | | | | | |
| **[2] ATT** | **0.729**$^{***}$ | **0.594**$^{***}$ | **0.578**$^{***}$ | **0.640**$^{***}$ | **0.621**$^{***}$ | **0.613**$^{***}$ | **0.620**$^{***}$ | **0.305**$^{***}$ |
| | (0.169) | (0.152) | (0.149) | (0.127) | (0.126) | (0.111) | (0.126) | (0.119) |
| Dependent variable: $Public\ Deficit_{i,t}$ | | | | | | | | |
| **[3] ATT** | **-2.217**$^{***}$ | **-1.990**$^{***}$ | **-2.074**$^{***}$ | **-2.155**$^{***}$ | **-2.154**$^{***}$ | **-2.147**$^{***}$ | **-2.139**$^{***}$ | **-1.160**$^{**}$ |
| | (0.678) | (0.529) | (0.644) | (0.616) | (0.412) | (0.421) | (0.504) | (0.360) |
| Dependent variable: $Debt\ growth\ rate_{i,t}$ | | | | | | | | |
| **[4] ATT** | **-2.922** | **-3.768** | **-3.382**$^{*}$ | **-5.064**$^{**}$ | **-4.576**$^{**}$ | **-4.973**$^{***}$ | **-4.549**$^{**}$ | **-4.755**$^{***}$ |
| | (3.216) | (2.726) | (2.491) | (2.599) | (2.536) | (1.901) | (2.055) | (1.521) |
| Dependent variable: $Interest\ growth\ rate_{i,t}$ | | | | | | | | |
| **[5] ATT** | **-5.164**$^{*}$ | **-2.931** | **-2.874** | **-3.098** | **-2.725** | **-3.205** | **-2.709** | **-2.212** |
| | (3.168) | (2.916) | (2.500) | (2.818) | (2.594) | (2.048) | (2.617) | (2.582) |
| Dependent variable: $Variations\ of\ fiscal\ revenues_{i,t}$ | | | | | | | | |
| **[6] ATT** | **-0.155** | **0.124** | **0.101** | **-0.074** | **-0.172** | **-0.191** | **-0.154** | **0.211** |
| | (0.529) | (0.448) | (0.484) | (0.412) | (0.488) | (0.450) | (0.506) | (0.427) |
| Dependent variable: $External\ Deficit_{i,t}$ | | | | | | | | |
| **[7] ATT** | **-27.948**$^{***}$ | **-30.076**$^{***}$ | **-27.416**$^{***}$ | **-30.586**$^{***}$ | **-29.962**$^{***}$ | **-30.613**$^{***}$ | **-30.256**$^{***}$ | **-23.852**$^{***}$ |
| | (7.365) | (6.969) | (7.305) | (7.021) | (6.576) | (6.346) | (6.416) | (4.419) |

Note: Bootstrapped standard errors (with 500 replications) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment. IPWRA includes all control variables for propensity scores estimation step except for ATT [6] and [7] where we removed the dependent variable from control variables included in propensity scores estimation.

Table 1.10: **Matching Results with DR (Debt Rules) as the treatment variable**

| | Nearest-neighbor Matching | | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Dependent variable: $CAPB_{i,t}$ | | | | | | | | |
| **[1] ATT** | **0.330** | **0.115** | **0.249** | **0.141** | **0.138** | **0.129** | **0.139** | **0.043** |
| | (0.460) | (0.361) | (0.350) | (0.365) | (0.253) | (0.300) | (0.247) | (0.204) |
| Number of treated observations | 90 | 90 | 88 | 90 | 90 | 90 | 90 | 90 |
| Number of control observations | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 |
| Standardized bias (p-value) | 0.914 | 0.936 | 0.845 | 0.923 | 0.931 | 0.914 | 0.939 | - |
| Dependent variable: $GFPI_{i,t}$ | | | | | | | | |
| **[2] ATT** | **0.136** | **0.153** | **0.157** | **0.100** | **0.148** | **0.145** | **0.136** | **0.099** |
| | (0.221) | (0.191) | (0.208) | (0.207) | (0.151) | (0.166) | (0.161) | (0.134) |
| Dependent variable: $Public\ Deficit_{i,t}$ | | | | | | | | |
| **[3] ATT** | **-1.014** | **-0.484** | **-0.866** | **-0.364** | **-0.427** | **-0.425** | **-0.407** | -0.497* |
| | (0.674) | (0.625) | (0.626) | (0.581) | (0.463) | (0.421) | (0.496) | (0.380) |
| Dependent variable: $Debt\ growth\ rate_{i,t}$ | | | | | | | | |
| **[4] ATT** | **-1.755** | **-0.122** | **-4.010** | **-0.764** | **-0.370** | **-0.474** | **-0.426** | **-0.193** |
| | (4.388) | (3.445) | (3.510) | (3.685) | (2.788) | (2.803) | (3.147) | (1.704) |
| Dependent variable: $Interest\ growth\ rate_{i,t}$ | | | | | | | | |
| **[5] ATT** | **2.227** | **1.213** | **-0.976** | **0.578** | **-0.130** | **0.024** | **0.237** | **1.433** |
| | (5.753) | (4.559) | (3.663) | (3.917) | (3.351) | (3.438) | (3.081) | (2.693) |
| Dependent variable: $Variations\ of\ fiscal\ revenues_{i,t}$ | | | | | | | | |
| **[6] ATT** | **0.290** | **0.086** | **-0.001** | **0.115** | **0.029** | **0.124** | **0.074** | **-0.008** |
| | (0.886) | (0.695) | (0.672) | (0.517) | (0.611) | (0.493) | (0.734) | (0.504) |
| Dependent variable: $External\ Deficit_{i,t}$ | | | | | | | | |
| **[7] ATT** | **-14.119*** | **-10.874*** | **-11.822** | **-10.204*** | **-9.812** | **-10.465*** | **-9.947** | **-10.972*** |
| | (7.982) | (5.967) | (8.517) | (7.298) | (7.464) | (6.320) | (5.646) | (4.189) |

Note: Bootstrapped standard errors (with 500 replications) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment. IPWRA includes all control variables for propensity scores estimation step except for ATT [6] and [7] where we removed the dependent variable from control variables included in propensity scores estimation.

## 1.7.2 The composition of the GFPI

We consider an alternative computation of the GFPI, by using the Public Expenditure Growth Rate[16]—instead of the Public Deficit—for the computation of the index.[17] The

[16]The corresponding primary indicator is Total Expenditures of Public Administrations in % of GDP from the Eurostat dataset "Government revenue, expenditure and main aggregates", and the level 2 index is the Public Expenditure Growth Rate calculated by authors and submitted to the standardization procedure.

[17]Unfortunately, since public deficit and the growth rate of public expenditure are correlated, they can not be used jointly to build the GFPI. In addition, as shown by the correlation matrix displayed in Appendix 13, this change does not raise double-counting issues between the primary indicators used to compute the

use of the growth rate, rather than of the level of public expenditure, is motivated by its common use for the monitoring medium-term public finance objectives.

Estimations reported in Table 11 show that the impact of fiscal rules on the GFPI is qualitatively unchanged when considering the growth rate of public expenditure instead of the public deficit. In a nutshell, although of a lower magnitude, FR, BBR and ER exert a positive and significant effect on GFPI, while DR has no significant effect. However, an additional heterogeneity arises in the impact of each type of fiscal rule on the new added variable, i.e. the growth rate of public expenditure: as shown by Table 12, while DR and—to some extent—ER significantly reduce the growth rate of public expenditure, the effect of BBR is mostly not significant. These findings confirm the robustness of our analysis with respect to the composition of the GFPI, and support the heterogeneity of the effect of various types of fiscal rules on the various variables that compose the GFPI index.

### 1.7.3 Structural factors

Having revealed that the effect of fiscal rules on fiscal discipline varies between the different types of fiscal rules, we now investigate if this effect may be subject to heterogeneity. To this end, we estimate the following control function regression

$$Y_{i,t} = \alpha + \beta FR_{i,t} + \gamma\, PS_{i,t} + \varphi X_{i,t} + \delta(FR_{i,t}X_{i,t}) + \epsilon_{i,t}, \tag{1.5}$$

with Y the measure of fiscal discipline (CAPB or GFPI), PS the propensity score that controls for reverse causality (see Rosenbaum and Rubin [1983]), and X the vector of factors that may trigger the heterogeneity in the effect of fiscal rules. We consider three groups of factors. First, macroeconomic factors include real GDP per capita, the lagged value of debt (in ratio of GDP), and two measures that capture difficult times. On the one hand, we use a binary variable (named "Bad times") equal to one during the years of financial crisis (2007-2008) and sovereign debt crisis (2010-2011). On the other hand, since this measure could be imperfect (see Sancak et al. [2010], and Boschi and d'Addona [2019]), we follow Wiese et al. [2018] and use a Bai-Perron test to identify structural breaks in the GFPI (reported in Appendix 11); as such, a decrease in the GFPI after the break signals a negative structural change in the fiscal behavior (we name this variable "Negative structural changes"). Second, political factors include the mode of election and—following e.g. Eklou and Joanis [2019] or Gootjes et al. [2019]—electoral political cycles. Third, fiscal-rule

new GFPI.

Table 1.11: **Matching Results on GFPI based on Public Expenditure Growth Rate**

| Dependent variable: $GFPI_{i,t}$ | Nearest-neighbor Matching | | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Treatment variable: FR | | | | | | | | |
| **[1] ATT** | **0.325**$^{**}$ | **0.320**$^{**}$ | **0.325**$^{**}$ | **0.271**$^{**}$ | **0.262**$^{**}$ | **0.274**$^{**}$ | **0.260**$^{**}$ | **0.155**$^{*}$ |
| | (0.170) | (0.150) | (0.144) | (0.133) | (0.135) | (0.127) | (0.132) | (0.111) |
| Number of treated observations | 203 | 203 | 191 | 203 | 203 | 203 | 203 | 203 |
| Number of control observations | 188 | 188 | 188 | 188 | 188 | 180 | 188 | 188 |
| Standardized bias (p-value) | 0.628 | 0.898 | 0.714 | 0.992 | 0.997 | 0.628 | 0.997 | - |
| Treatment variable: BBR | | | | | | | | |
| **[2] ATT** | **0.321**$^{*}$ | **0.321**$^{**}$ | **0.296**$^{**}$ | **0.298**$^{**}$ | **0.282**$^{**}$ | **0.258**$^{**}$ | **0.278**$^{**}$ | **0.197**$^{*}$ |
| | (0.189) | (0.166) | (0.158) | (0.156) | (0.147) | (0.131) | (0.141) | (0.128) |
| Number of treated observations | 108 | 108 | 104 | 108 | 108 | 108 | 108 | 116 |
| Number of control observations | 276 | 276 | 276 | 276 | 276 | 276 | 276 | 276 |
| Standardized bias (p-value) | 0.919 | 0.796 | 0.974 | 0.935 | 0.961 | 0.919 | 0.958 | - |
| Treatment variable: ER | | | | | | | | |
| **[3] ATT** | **0.488**$^{***}$ | **0.401**$^{***}$ | **0.357**$^{***}$ | **0.429**$^{***}$ | **0.437**$^{***}$ | **0.440**$^{***}$ | **0.435**$^{***}$ | **0.251**$^{**}$ |
| | (0.166) | (0.155) | (0.143) | (0.137) | (0.122) | (0.130) | (0.130) | (0.117) |
| Number of treated observations | 121 | 121 | 117 | 120 | 121 | 121 | 121 | 122 |
| Number of control observations | 270 | 270 | 270 | 270 | 270 | 270 | 270 | 270 |
| Standardized bias (p-value) | 0.898 | 0.977 | 0.953 | 0.954 | 0.989 | 0.898 | 0.987 | - |
| Treatment variable: DR | | | | | | | | |
| **[4] ATT** | **0.198** | **0.273** | **0.242** | **0.206** | **0.223** | **0.214** | **0.213** | **0.205** |
| | (0.240) | (0.205) | (0.213) | (0.179) | (0.185) | (0.180) | (0.185) | (0.150) |
| Number of treated observations | 90 | 90 | 88 | 90 | 90 | 90 | 90 | 90 |
| Number of control observations | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 |
| Standardized bias (p-value) | 0.914 | 0.936 | 0.845 | 0.923 | 0.931 | 0.914 | 0.939 | - |

Note: Bootstrapped standard errors (with 500 replications) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment and uses saturated propensity scores.

related factors include the number of years during which a national rule has been in force, the presence of the Stability and Growth Pact (SGP), the presence of an independent institution in charge of the fiscal discipline monitoring, and the number of rules in force.

Results are reported in Table 13 (for the CAPB) and Table 14 (for the GFPI). In particular, the significance of the coefficient of the propensity score supports—once again—the importance of controlling for the self-selection bias by using the propensity score matching method. The effect of the different variables can be classified in three groups. First, out of the ten variables considered, two of them exert the same type of significant effect on the two measures of fiscal discipline, namely, CAPB and GFPI. A higher (lagged) public debt ratio reduces the favorable effect of fiscal rules on the CAPB and the GFPI, probably

Table 1.12: **Matching Results on Public Expenditure Growth Rate**

| Dependent variable: | Nearest-neighbor Matching | | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|
| $Public\ Expenditure\ Growth\ Rate_{i,t}$ | N = 1 | N = 3 | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Treatment variable: FR | | | | | | | | |
| **[1] ATT** | **-1.355*** | **-1.015**** | **-1.560**** | **-1.225*** | **-1.227*** | **-1.424**** | **-1.196*** | **-1.689***** |
| | (1.049) | (0.897) | (0.822) | (0.838) | (0.734) | (0.751) | (0.717) | (0.672) |
| Number of treated observations | 203 | 203 | 191 | 203 | 203 | 203 | 203 | 203 |
| Number of control observations | 188 | 188 | 188 | 188 | 188 | 180 | 188 | 188 |
| Standardized bias (p-value) | 0.628 | 0.898 | 0.714 | 0.992 | 0.997 | 0.628 | 0.997 | - |
| Treatment variable: BBR | | | | | | | | |
| **[2] ATT** | **-1.663*** | **-1.147** | **-0.921**** | **-1.089** | **-1.050** | **-1.020** | **-1.009** | **-1.331*** |
| | (1.151) | (1.023) | (0.959) | (0.903) | (0.848) | (0.827) | (0.825) | (0.826) |
| Number of treated observations | 108 | 108 | 104 | 108 | 108 | 108 | 108 | 116 |
| Number of control observations | 276 | 276 | 276 | 276 | 276 | 276 | 276 | 276 |
| Standardized bias (p-value) | 0.919 | 0.796 | 0.974 | 0.935 | 0.961 | 0.919 | 0.958 | - |
| Treatment variable: ER | | | | | | | | |
| **[3] ATT** | **-0.488** | **-1.328** | **-1.011** | **-1.484**** | **-1.589**** | **-1.685**** | **-1.561**** | **-1.870***** |
| | (1.068) | (0.869) | (0.922) | (0.742) | (0.747) | (0.742) | (0.784) | (0.724) |
| Number of treated observations | 121 | 121 | 117 | 120 | 121 | 121 | 121 | 122 |
| Number of control observations | 270 | 270 | 270 | 270 | 270 | 270 | 270 | 270 |
| Standardized bias (p-value) | 0.898 | 0.977 | 0.953 | 0.954 | 0.989 | 0.898 | 0.987 | - |
| Treatment variable: DR | | | | | | | | |
| **[4] ATT** | **-2.560*** | **-1.790*** | **-2.480**** | **-1.786**** | **-1.353*** | **-1.397*** | **-1.412*** | **-1.360*** |
| | (1.368) | (1.045) | (1.213) | (0.909) | (0.771) | (0.791) | (0.829) | (0.738) |
| Number of treated observations | 90 | 90 | 88 | 90 | 90 | 90 | 90 | 90 |
| Number of control observations | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 |
| Standardized bias (p-value) | 0.914 | 0.936 | 0.845 | 0.923 | 0.931 | 0.914 | 0.939 | - |

Note: Bootstrapped standard errors (with 500 replications) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment and uses saturated propensity scores.

due to a crowding-out effect of a larger debt burden in the presence of high indebtedness. Conversely, the presence of an electoral system characterized by a president elected by assembly (against directly-elected) or by a parliamentary system (against a president elected by assembly, or a directly-elected president) significantly improves the favorable effect of fiscal rules on both CAPB and the GFPI.[18] Second, some variables do not significantly affect the impact of fiscal rules on fiscal discipline irrespective of its measure, namely, the (log of) real GDP per capita, electoral cycles, and the presence of monitoring institutions. Third, some variables significantly influence the effect of fiscal rules on the

[18] These findings may be related to the conclusions of Hallerberg et al. [2007], emphasizing the importance of the forms of governance for fiscal discipline.

CAPB but not the GFPI, and conversely. In the former group, bad times and structural negative changes (the number of years covered by the rule) decrease (increases) only the favorable effect of fiscal rules on the CAPB—and do not affect fiscal discipline measured by the GFPI. In the latter group, the presence of the SGP and of a larger number of fiscal rules positively influences the effect of fiscal rules on the GFPI (but not on the CAPB). Particularly regarding the SGP, its significant contribution to fiscal discipline measured by the GFPI may contribute to the current debate on the various propositions of reform it (see e.g. Darvas et al. [2018] or Hauptmeier and Kamps [2020]).

Altogether, these results show that the favorable effect of fiscal rules on fiscal discipline may be altered by various factors that seize different structural characteristics. Corroborating our previous findings, they confirm that the effect of such factors is fairly different when using alternative measures of fiscal discipline.[19]

[19]Similar conclusions are found when using a logit, instead of a probit model, to compute propensity scores (results are available upon request).

Table 1.13: **Nonlinearities in the effect of FR on the CAPB**

| Variables | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] | [9] | [10] |
|---|---|---|---|---|---|---|---|---|---|---|
| Dummy variable FR | 0.006 | 1.022** | 0.584*** | 0.600*** | −0.260 | 0.651** | 0.232 | 0.553** | 0.528** | 0.357*** |
| | (1.473) | (0.364) | (0.226) | (0.232) | (0.429) | (0.271) | (0.438) | (0.239) | (0.250) | (0.347) |
| Propensity Score | −1.515*** | −1.384** | −1.547*** | −1.526*** | −1.529*** | −1.534*** | −1.548*** | −1.603*** | −1.555*** | −0.362*** |
| | (0.565) | (0.724) | (0.519) | (0.552) | (0.518) | (0.551) | (0.556) | (0.612) | (0.579) | (0.563) |
| **Macroeconomics Factors** | | | | | | | | | | |
| FR * log Real gdp per capita | 0.051 | | | | | | | | | |
| | (0.144) | | | | | | | | | |
| FR * Debt ratio $_{t-1}$ | | −0.009* | | | | | | | | |
| | | (0.006) | | | | | | | | |
| FR * Bad Time | | | −0.118* | | | | | | | |
| | | | (0.553) | | | | | | | |
| FR * Negative Structural Changes | | | | −0.834* | | | | | | |
| | | | | (0.567) | | | | | | |
| **Political factors** | | | | | | | | | | |
| FR * Electoral system | | | | | 0.882** | | | | | |
| | | | | | (0.473) | | | | | |
| FR * Electoral cycles | | | | | | −0.316 | | | | |
| | | | | | | (0.510) | | | | |
| **Factors linked with Rules** | | | | | | | | | | |
| FR * Number years covered by rules | | | | | | | 0.095* | | | |
| | | | | | | | (0.071) | | | |
| FR * SGP | | | | | | | | −0.008 | | |
| | | | | | | | | (0.374) | | |
| FR * Monitoring institution | | | | | | | | | −0.191 | |
| | | | | | | | | | (0.390) | |
| FR * Number of rules | | | | | | | | | | 0.108 |
| | | | | | | | | | | (0.150) |
| Observations | 392 | 392 | 392 | 392 | 392 | 392 | 392 | 392 | 392 | 392 |
| p-value Chi2 test | 0.03 | 0.011 | 0.016 | 0.010 | 0.011 | 0.012 | 0.012 | 0.03 | 0.00 | 0.01 |

Note: FGLS estimator is used. Bootstrapped standard errors (with 500 replications based on clustering on country level) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. For each column, the intercept and the variable not interacted with FR are included but not reported.

Table 1.14: **Nonlinearities in the effect of FR on the GFPI**

| Variables | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] | [9] | [10] |
|---|---|---|---|---|---|---|---|---|---|---|
| Dummy variable FR | 0.511*** | 1.025*** | 0.447*** | 0.468*** | −0.607** | 0.550** | 0.836*** | 0.290* | 0.413*** | −0.070 |
| | (0.105) | (0.231) | (0.132) | (0.329) | (0.295) | (0.127) | (0.316) | (0.156) | (0.128) | (0.198) |
| Propensity Score | −0.429* | −0.207 | −0.367* | −0.401* | −0.247 | −0.378* | −0.270** | −0.223 | −0.125 | −0.436*** |
| | (0.257) | (0.331) | (0.271) | (0.268) | (0.288) | (0.245) | (0.129) | (0.264) | (0.277) | (0.267) |
| **Macroeconomics Factors** | | | | | | | | | | |
| FR * Real gdp per capita | −0.0000002 | | | | | | | | | |
| | (0.0000003) | | | | | | | | | |
| FR * Debt ratio$_{t-1}$ | | −0.011*** | | | | | | | | |
| | | (0.004) | | | | | | | | |
| FR * Bad Time | | | 0.098 | | | | | | | |
| | | | (0.208) | | | | | | | |
| FR * Negative Structural changes | | | | −0.005 | | | | | | |
| | | | | (0.514) | | | | | | |
| **Political factors** | | | | | | | | | | |
| FR * Electoral system | | | | | 1.239*** | | | | | |
| | | | | | (0.307) | | | | | |
| FR * Electoral cycles | | | | | | −0.232 | | | | |
| | | | | | | (0.204) | | | | |
| **Factors linked with Rules** | | | | | | | | | | |
| FR * Number years | | | | | | | −0.058 | | | |
| covered by rules | | | | | | | (0.043) | | | |
| FR * SGP | | | | | | | | 0.365** | | |
| | | | | | | | | (0.190) | | |
| FR * Monitoring institution | | | | | | | | | −0.099 | |
| | | | | | | | | | (0.115) | |
| FR * Number of rules | | | | | | | | | | 0.308*** |
| | | | | | | | | | | (0.090) |
| Observations | 392 | 392 | 392 | 392 | 392 | 392 | 392 | 392 | 392 | 392 |
| p-value Chi2 test | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

Note: FGLS estimator is used. Bootstrapped standard errors (with 500 replications based on clustering on country level) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. For each column, the intercept and the variable not interacted with FR are included but not reported.

## 1.8 Conclusion

Motivated by the burning debate on fiscal rules lack of performance followed by propositions to replace and/or reform them (Blanchard et al. [2021], Hauptmeier and Kamps [2020]), this chapter analyzed the effect of national fiscal rules on fiscal discipline, using a careful definition of national fiscal rules combined with a novel measure of fiscal discipline (namely, the Global Financial Performance Index—GFPI). Propensity score matching estimations that account for potential endogeneity revealed that the fiscal rules significantly improve the GFPI, corroborating their favorable effect on the popular measure of fiscal discipline—the CAPB—emphasized by some of the existing studies. Such conclusions help in answering to the skepticism about the usefulness and effectiveness of fiscal rules. Moreover, the results are robust to various alternative specifications but they dramatically depend on the type of fiscal rule and different structural factors (i.e. countries' and rules' structural characteristics). Together with alternative fiscal discipline measures, these features must be taken into account when assessing the effects of fiscal rules on fiscal discipline. This chapter therefor allows to see that some fiscal rules are more effective than others and should be considered in the economic area to which they belong. These results are more important with regard to inter-country heterogeneities, and even more so in the monetary unions.

We see two first-order policy implications of this chapter. First, it is of particular importance to use different measures of fiscal discipline when assessing its response to the presence of various types of fiscal rules, since the effects of fiscal rules may dramatically differ both in significance and magnitude. Second, when following a fiscal discipline goal, it would be of interest to imagine fiscal rules that may account for variations in structural factors (i.e. countries' and rules' characteristics), since such factors can boost, or—on the contrary—mitigate the favorable effects of fiscal rules.

This chapter finally open the door to future work. First, close to our study, it would be interesting to look at the response of fiscal discipline to the so-called second-generation fiscal rules (see Eyraud et al. [2018]), which potentially add flexibility and enforceability to the simplicity feature of the traditional fiscal rules—see the discussions relative to the "fiscal rules trilemma" in e.g. Debrun and Jonung [2019]. Second, beyond national fiscal rules, one could explore the relationship between sub-national fiscal rules and fiscal discipline, from a cross-country perspective. Third, since our empirical analysis was conducted on EU countries, future studies could investigate the nature of the effect of fiscal rules on fiscal discipline in other economic and monetary areas, including the two African monetary unions—the CEMAC and the WAEMU.

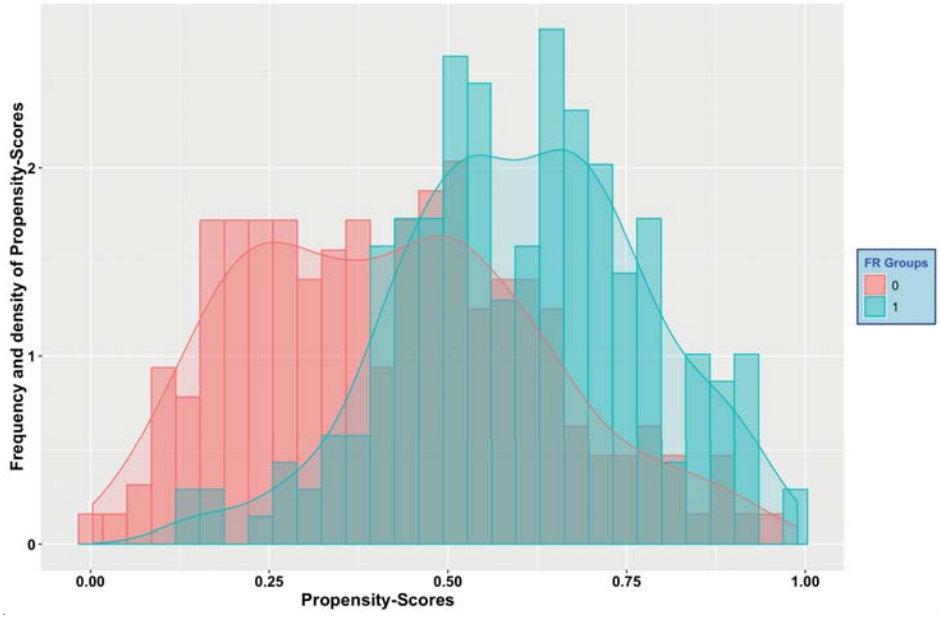
# Appendices

## Appendix 1. National numerical fiscal rules excluded by our definition

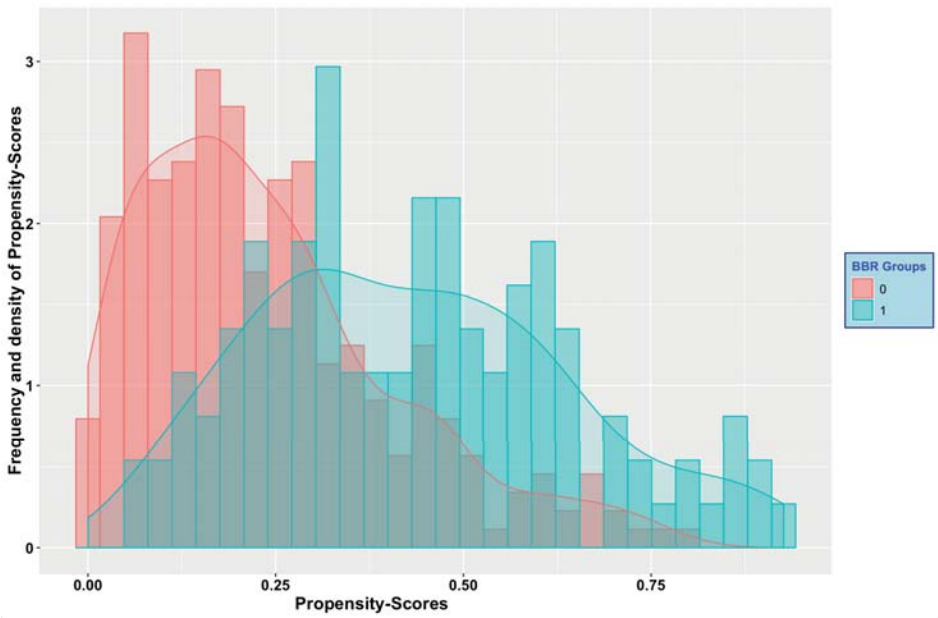
| Countries/Fiscal Rules Excluded | BBR | DR | ER | RR |
|---|---|---|---|---|
| Austria | 2000-2013: MTBF<br>(IMF Fiscal Rules Database and Reuter, 2015) | | | |
| Belgium | Belgium adopted a BBR in 2014 (according to IMF and European Commission databases), so it does not have a fiscal rule during our study period | | | |
| France | 2013: MTBF. The rule is written in the public finance programming law that can be revised, so it is not comparable with a numerical fiscal rule described by Kopits and Symansky (1998) | | | 2006-2013: MTBF |
| United Kingdom | 2009: Fiscal rule abandoned during 2009 (IMF fiscal rules database and Reuter, 2015) | 2009: Fiscal rule abandoned during 2009 (IMF fiscal rules database).<br>2010: Fiscal rule also abandoned in 2010. | | |

Note: MTBF stands for Medium Term Budgetary Framework.

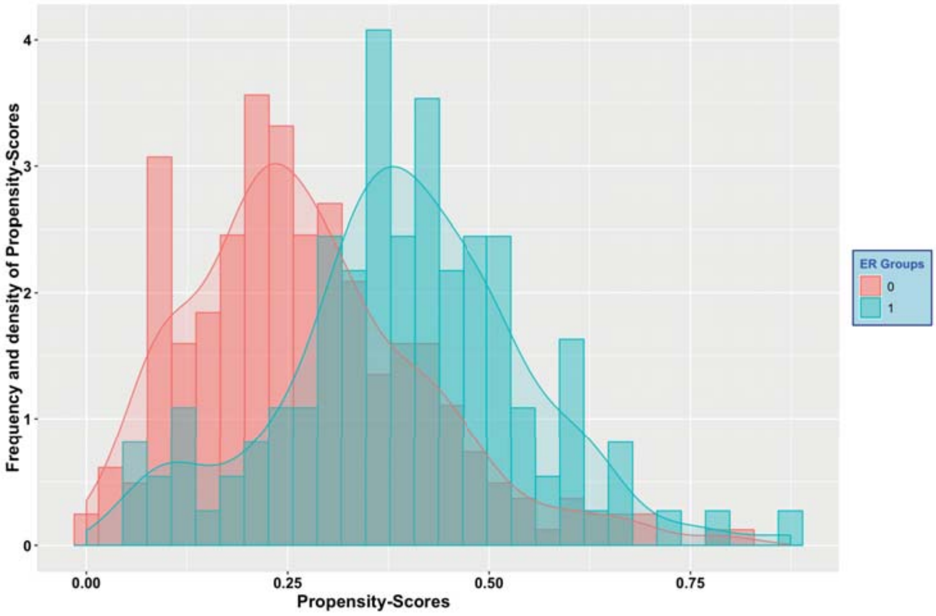
**Appendix 2.1: Common Support Region for FR**



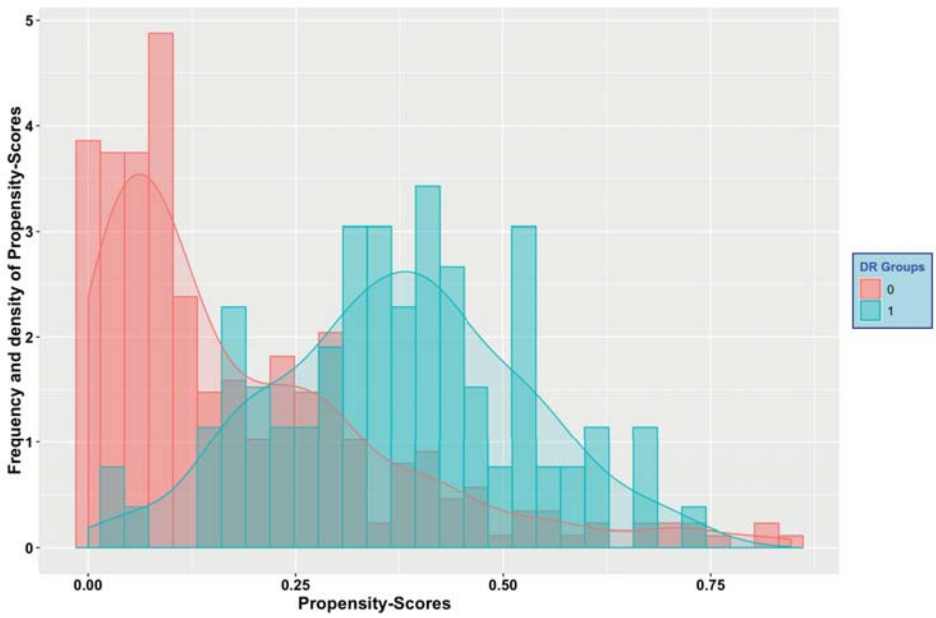

**Appendix 2.2: Common Support Region for BBR**



**Appendix 2.3: Common Support Region for ER**




**Appendix 2.4: Common Support Region for DR**



## Appendix 3: Correlations between the four indicators used to construct the GFPI

| | Total Budget Balance | External Deficit | Variations of Fiscal Revenues | Sustainability Debt Index |
|---|---|---|---|---|
| Total Budget Balance | 1.000 | -0.338 | -0.010 | -0.099 |
| External Deficit | - | 1.000 | -0.002 | 0.052 |
| Variations of Fiscal Revenues | - | - | 1.000 | 0.017 |
| Sustainability Debt Index | - | - | - | 1.000 |

## Appendix 4. Probit estimates of the Propensity Scores—Robustness

| Dependent variable | [1]<br>FR | [2]<br>FR | [3]<br>FR | [4]<br>FR | [5]<br>FR | [6]<br>FR | [7]<br>FR | [8]<br>FR | [9]<br>FR | [10]<br>FR | [11]<br>FR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Intercept | -0.983 | -1.025 | -0.966 | -0.962 | -1.053 | -0.747 | -1.100 | -0.934 | -0.992 | -0.769 | -0.919 |
| | (1.048) | (1.048) | (1.031) | (1.058) | (1.063) | (1.058) | (1.058) | (1.054) | (1.067) | (1.064) | (3.868) |
| CAPB$_{t-1}$ | 0.113*** | 0.119*** | 0.118*** | 0.115*** | 0.113*** | 0.108*** | 0.109*** | 0.114*** | 0.113*** | 0.110*** | 0.137** |
| | (0.042) | (0.043) | (0.043) | (0.043) | (0.043) | (0.043) | (0.043) | (0.042) | (0.042) | (0.042) | (0.072) |
| Debt ratio$_{t-1}$ | -0.018*** | -0.017*** | -0.019*** | -0.019*** | -0.018* | -0.026*** | -0.018*** | -0.018*** | -0.018*** | -0.019*** | -0.043*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.006) |
| Real per capita GDP growth rate | -0.028 | -0.025 | -0.033 | -0.026 | -0.029 | -0.027 | -0.029 | -0.029 | -0.027 | -0.030 | -0.078 |
| | (0.021) | (0.021) | (0.020) | (0.020) | (0.020) | (0.020) | (0.020) | (0.020) | (0.020) | (0.020) | (0.056) |
| Inflation rate | -0.103*** | -0.130* | -0.099*** | -0.102*** | -0.104*** | -0.105*** | -0.103*** | -0.100*** | -0.104*** | -0.098*** | -0.167* |
| | (0.026) | (0.025) | (0.025) | (0.026) | (0.027) | (0.026) | (0.026) | (0.026) | (0.027) | (0.027) | (0.117) |
| Government stability | 0.065* | 0.046 | 0.016 | 0.073 | 0.076 | 0.137 | 0.037 | 0.112 | 0.066 | -0.082 | -0.199 |
| | (0.200) | (0.198) | (0.200) | (0.201) | (0.204) | (0.203) | (0.201) | (0.201) | (0.200) | (0.201) | (0.393) |
| SGP | -0.080 | -0.094 | -0.104 | -0.072 | -0.086 | -0.095 | -0.036 | -0.036 | -0.074 | -0.078 | -1.522*** |
| | (0.162) | (0.161) | (0.163) | (0.162) | (0.164) | (0.162) | (0.164) | (0.164) | (0.177) | (0.162) | (0.494) |
| Dummy EU membership | 0.077 | 0.048 | 0.068 | 0.058 | 0.086 | 0.091 | 0.077 | -0.077 | 0.076 | 0.118 | - |
| | (0.386) | (0.381) | (0.379) | (0.388) | (0.388) | (0.387) | (.390) | (0.390) | (0.387) | (1.068) | - |
| Unemployment rate | 0.030* | 0.028* | 0.050** | 0.030* | 0.033* | 0.027* | 0.028* | 0.029* | 0.030* | 0.027 | 0.075* |
| | (0.019) | (0.019) | (0.020) | (0.019) | (0.022) | (0.019) | (0.019) | (0.019) | (0.019) | (0.019) | (0.040) |
| REER | 0.026*** | 0.026*** | 0.027*** | 0.025*** | 0.026*** | 0.026*** | 0.027*** | 0.028*** | 0.026*** | 0.024*** | 0.054* |
| | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.036) |
| Trade openness | -0.008** | -0.007** | -0.009*** | -0.008** | -0.008*** | -0.008** | -0.008** | -0.008*** | -0.008*** | -0.009*** | -0.010** |
| | (0.003) | (0.003) | (0.002) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.005) |
| SFA on debt | | -0.037* | | | | | | | | | |
| | | (0.022) | | | | | | | | | |
| Adding external deficit | | | -0.005** | | | | | | | | |
| | | | (0.002) | | | | | | | | |
| Adding Variations of Fiscal Revenues | | | | -0.015 | | | | | | | |
| | | | | (0.023) | | | | | | | |
| Adding output gap | | | | | 1.410 | | | | | | |
| | | | | | (5.347) | | | | | | |
| Adding lagged squared debt | | | | | | 0.00006 | | | | | |
| | | | | | | (0.00006) | | | | | |
| Adding gov. fragmentation | | | | | | | 0.266 | | | | |
| | | | | | | | (0.303) | | | | |
| Electoral system | | | | | | | | 0.189* | | | |
| | | | | | | | | (0.113) | | | |
| Emerging country | | | | | | | | | 0.017 | | |
| | | | | | | | | | (0.197) | | |
| PSC reforms | | | | | | | | | | 0.288* | |
| | | | | | | | | | | (0.172) | |
| Excl. New EU & Greece | | | | | | | | | | | |
| Adjusted $R^2$ | 0.097 | 0.142 | 0.106 | 0.094 | 0.093 | 0.095 | 0.095 | 0.098 | 0.093 | 0.098 | 0.415 |
| Observations | 392 | 392 | 392 | 392 | 392 | 392 | 392 | 392 | 392 | 392 | 196 |

Note: Robust standard errors in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. In column [10] the probit is estimated when excluding the new EU countries (that entered the EU after 2004) and Greece (since all remaining countries were in EU, the dummy EU membership is dropped).

### Appendix 5. Probit estimates of the Propensity Scores for BBR, ER, and DR

| Dependent variable | BBR | ER | DR |
|---|---|---|---|
| Intercept | -1.995* | -0.157 | -4.259*** |
| | (1.154) | (1.188) | (1.179) |
| $CAPB_{t-1}$ | 0.110** | 0.070* | 0.037 |
| | (0.048) | (0.042) | (0.047) |
| Debt $ratio_{t-1}$ | -0.015*** | -0.007** | -0.021*** |
| | (0.003) | (0.003) | (0.004) |
| Real per capita GDP growth rate | -0.018 | -0.074*** | 0.015 |
| | (0.020) | (0.020) | (0.020) |
| Inflation rate | -0.091*** | -0.062* | -0.063*** |
| | (0.027) | (0.037) | (0.024) |
| Government stability | 0.391* | 0.552*** | 0.609*** |
| | (0.207) | (0.169) | (0.241) |
| SGP | -0.379** | 0.201 | -0.401** |
| | (0.168) | (0.165) | (0.194) |
| Dummy EU membership | 0.089 | -0.055 | 0.299 |
| | (0.394) | (0.446) | (0.439) |
| Unemployment rate | 0.016 | -0.0006 | 0.073*** |
| | (0.021) | (0.020) | (0.021) |
| REER | 0.033*** | 0.0005 | 0.036*** |
| | (0.010) | (0.010) | (0.009) |
| Trade openness | -0.019*** | -0.004 | 0.001 |
| | (0.004) | (0.003) | (0.004) |
| Adjusted $R^2$ | 0.170 | 0.105 | 0.200 |
| Observations | 392 | 392 | 392 |

Note: Robust standard errors in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively.

## Appendix 6. Sources of all the variables used in the study

| *Variable* | *Source* |
|---|---|
| Debt/GDP ratio | IMF Historical Database |
| Term of trade (index) | IMF |
| Primary Balance | AMECO Database |
| Revenues of public administrations | Eurostat |
| Expenditure of public administrations | Eurostat |
| Inflation | IMF |
| Commodity Price Index | Federal Reserve Bank of St Louis |
| Real per capita GDP growth rate | World Bank |
| Population | World Bank |
| Government stability | World Bank (WGI) |
| Dependency ratio | World Bank (WGI) |
| Government fragmentation | World Bank (DPI 2015) |
| Electoral system | World Bank (DPI 2015) |
| Electoral cycles | World Bank (DPI 2015) |
| External deficit | Eurostat |
| Fiscal rules | IMF Fiscal Rules Database |
| Number of rules | IMF Fiscal Rules Database |
| Number of years covered by rules | Authors' calculations |
| Total budget balance | IMF |
| Cyclically adjusted balance (Hodrick Prescott filter) | Authors' calculations |
| Cyclically adjusted balance (Trigonometric filter) | Authors' calculations |
| Cyclically adjusted balance (production function approach) | IMF |
| Interest on debt | World Bank (WDI) |
| Output gap (Hodrick Prescott filter) | Authors' calculations |
| Output gap (Trigonometric filter) | Authors' calculations |
| Real effective exchange rate | Eurostat |
| Trade openness | OECD |
| Stock-flow adjustment on gen. gov. consolidated gross debt | AMECO Database |
| Real GDP | Eurostat |
| Stock Flow adjustment on debt | Authors' calculations |
| Negative Structural Changes in GFPI | Authors' calculations |

## Appendix 7. Descriptive statistics

| Variable | N | Mean | Min | Max | sd |
|---|---|---|---|---|---|
| Public debt (% of GDP) | 392 | 53.202 | 3.664 | 177.677 | 30.37 |
| Term of trade (index) | 392 | 0.9975 | 0.8906 | 1.2320 | 0.0398 |
| Inflation | 392 | 3.27 | -1.70 | 45.70 | 3.6541 |
| Commodity price index | 392 | 120.63 | 58.25 | 192.57 | 49.013 |
| Real per capita GDP growth rate | 392 | 2.053 | -14.559 | 12.920 | 3.936 |
| Government stability | 392 | 0.8045 | -0.7798 | 1.7602 | 0.4635 |
| Government fragmentation | 392 | 0.3716 | 0.0000 | 0.8278 | 0.2582 |
| Electoral cycles | 392 | 0.3214 | 0.0000 | 1.0000 | 0.4676 |
| Electoral system | 392 | 1.735 | 0.000 | 2.000 | 0.6484 |
| External deficit | 392 | 30.76 | -140.30 | 156.00 | 44.3296 |
| Fiscal rules | 392 | 0.5204 | 0.0000 | 1.0000 | 0.5002 |
| Expenditure rules | 392 | 0.2959 | 0.0000 | 1.0000 | 0.4636 |
| Budget balance rules | 392 | 0.2959 | 0.0000 | 1.0000 | 0.4570 |
| Debt rules | 392 | 0.2296 | 0.0000 | 1.0000 | 0.4211 |
| PSC reforms | 392 | 0.2143 | 0.0000 | 1.0000 | 0.4108 |
| Number of national fiscal rules | 392 | 0.9388 | 0.0000 | 3.0000 | 1.0346 |
| Number of years covered by fiscal rules | 392 | 7.158 | 0.000 | 14.000 | 6.0602 |
| Total budget balance (% of GDP) | 392 | -2.794 | -32.000 | 6.700 | 3.7569 |
| Cyclically adjusted primary balance - Hodrick Prescott filter (% of GDP) | 392 | 0.0000 | -19.744 | 11.076 | 1.950 |
| Cyclically adjusted primary balance - Trigonometric filter (% of GDP) | 392 | 0.0000 | -19.552 | 11.171 | 1.9629 |
| Cyclically adjusted primary balance - IMF production function approach (% of GDP) | 356 | -0.9007 | -10.672 | 7.8373 | 2.9439 |
| Global fiscal performance index (GFPI) (% of GDP) | 392 | 0.0000 | -2.2001 | 4.6229 | 1.0000 |
| Growth of debt interest | 392 | 4.166 | -56.075 | 126.05 | 17.685 |
| Output gap (Hodrick Prescott filter) | 392 | -0.009 | -4.7102 | 7.5252 | 1.5986 |
| Real Effective Exchange Rate | 392 | 98.51 | 66.07 | 184.36 | 9.5713 |
| Trade openness | 392 | 55.83 | 22.23 | 142.63 | 24.725 |
| Emerging country | 392 | 0.2143 | 0.0000 | 1.0000 | 0.4108 |
| Variations of Fiscal Revenues (with a negative sign) | 392 | -0.1746 | -18.329 | 10.488 | 3.117 |
| Public Expenditure Growth Rate | 392 | 0.4395 | -28.2642 | 38.5106 | 5.4906 |
| Dummy EU membership | 392 | 0.0000 | 0.0332 | 1.0000 | 0.1793 |
| Lagged squared debt ratio | 392 | 3518.5 | 13.42 | 29617.1 | 3855.7 |
| Unemployment rate | 392 | 8.819 | 1.805 | 27.466 | 4.2969 |
| Stock flow adjustment on debt (% of GDP) | 392 | 0.68 | -35.61 | 13.79 | 3.516 |
| Negative structural changes in GFPI | 392 | 0.061 | 0.000 | 1.000 | 2.240 |

## Appendix 8: Matching Results on CAPB (country based cluster-boostrapped errors)

| Dependent variable: $CAPB_{i,t}$ | Nearest-neighbor Matching | | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Treatment variable: FR | | | | | | | | |
| **[1] ATT** | **0.698**** | **0.451** | **0.676**** | **0.540*** | **0.515**** | **0.517**** | **0.510*** | **0.290**** |
| | (0.409) | (0.335) | (0.339) | (0.342) | (0.263) | (0.288) | (0.326) | (0.150) |
| Number of treated observations | 203 | 203 | 191 | 203 | 203 | 203 | 203 | 203 |
| Number of control observations | 188 | 188 | 188 | 188 | 188 | 180 | 188 | 188 |
| Standardized bias (p-value) | 0.628 | 0.898 | 0.714 | 0.992 | 0.997 | 0.628 | 0.997 | - |
| Treatment variable: BBR | | | | | | | | |
| **[2] ATT** | **0.297** | **0.465**** | **0.423**** | **0.436*** | **0.502**** | **0.494**** | **0.509**** | **0.344**** |
| | (0.258) | (0.335) | (0.314) | (0.262) | (0.228) | (0.231) | (0.212) | (0.166) |
| Number of treated observations | 108 | 108 | 104 | 108 | 108 | 108 | 108 | 116 |
| Number of control observations | 276 | 276 | 276 | 276 | 276 | 276 | 276 | 276 |
| Standardized bias (p-value) | 0.919 | 0.796 | 0.974 | 0.935 | 0.961 | 0.919 | 0.958 | - |
| Treatment variable: ER | | | | | | | | |
| **[3] ATT** | **0.571** | **0.267** | **0.264** | **0.284** | **0.395** | **0.371** | **0.386** | **0.152*** |
| | (0.520) | (0.399) | (0.288) | (0.405) | (0.294) | (0.403) | (0.284) | (0.163) |
| Number of treated observations | 121 | 121 | 117 | 120 | 121 | 121 | 121 | 122 |
| Number of control observations | 270 | 270 | 270 | 270 | 270 | 270 | 270 | 270 |
| Standardized bias (p-value) | 0.898 | 0.977 | 0.953 | 0.954 | 0.989 | 0.898 | 0.987 | - |
| Treatment variable: DR | | | | | | | | |
| **[4] ATT** | **0.330** | **0.115** | **0.249** | **0.141** | **0.138** | **0.130** | **0.139** | **0.043** |
| | (0.669) | (0.350) | (0.390) | (0.357) | (0.361) | (0.365) | (0.363) | (0.205) |
| Number of treated observations | 90 | 90 | 88 | 90 | 90 | 90 | 90 | 90 |
| Number of control observations | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 |
| Standardized bias (p-value) | 0.914 | 0.936 | 0.845 | 0.923 | 0.931 | 0.914 | 0.939 | - |

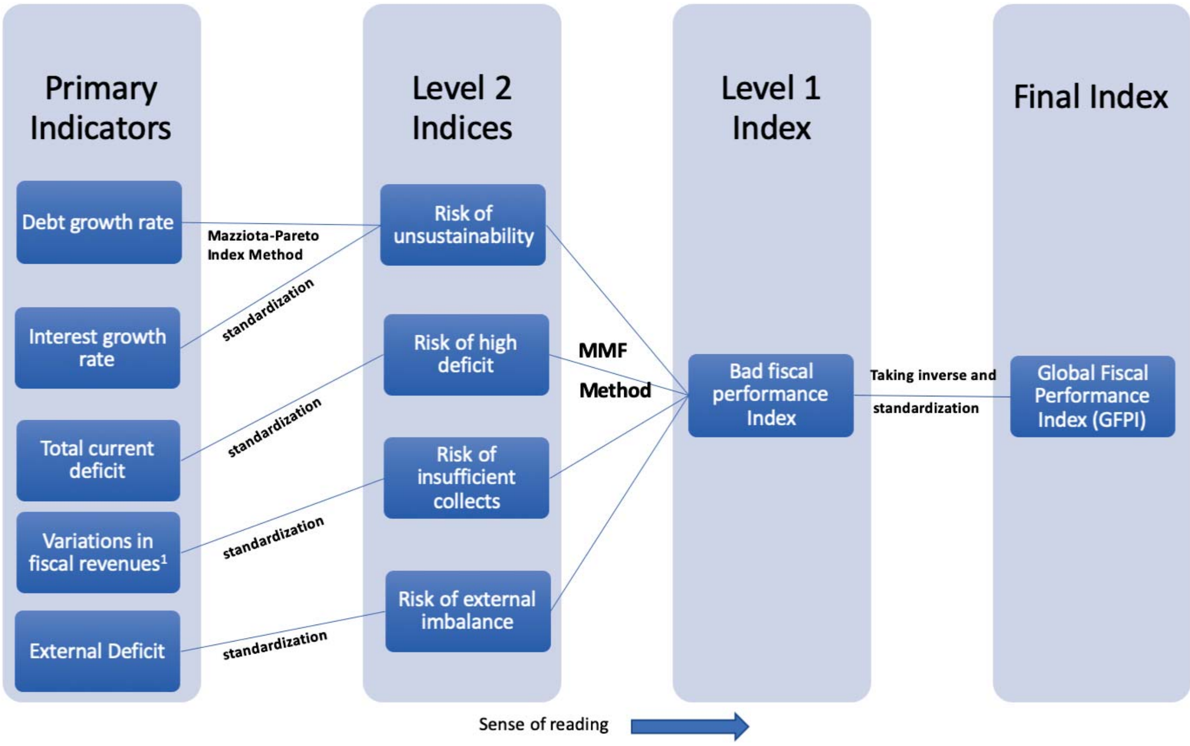
Note: country based cluster-bootstrapped standard errors in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment and uses saturated propensity scores.

## Appendix 9: Matching Results on GFPI (country based cluster-boostrapped errors)

| Dependent variable: $GFPI_{i,t}$ | Nearest-neighbor Matching | | Radius Matching | | | local linear Matching | kernel Matching | IPWRA |
|---|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | | |
| Treatment variable: FR | | | | | | | | |
| **[1] ATT** | **0.494**$^{**}$ | **0.537**$^{**}$ | **0.524**$^{**}$ | **0.504**$^{**}$ | **0.490**$^{**}$ | **0.499**$^{***}$ | **0.491**$^{***}$ | **0.351** |
| | (0.245) | (0.251) | (0.245) | (0.232) | (0.218) | (0.179) | (0.183) | (0.105) |
| Number of treated observations | 203 | 203 | 191 | 203 | 203 | 203 | 203 | 203 |
| Number of control observations | 188 | 188 | 188 | 188 | 188 | 180 | 188 | 188 |
| Standardized bias (p-value) | 0.628 | 0.898 | 0.714 | 0.992 | 0.997 | 0.628 | 0.997 | - |
| Treatment variable: BBR | | | | | | | | |
| **[2] ATT** | **0.679**$^{***}$ | **0.715**$^{***}$ | **0.703**$^{***}$ | **0.717**$^{**}$ | **0.702**$^{***}$ | **0.676**$^{***}$ | **0.673**$^{***}$ | **0.698**$^{***}$ |
| | (0.215) | (0.191) | (0.245) | (0.197) | (0.249) | (0.188) | (0.247) | (0.108) |
| Number of treated observations | 108 | 108 | 104 | 108 | 108 | 108 | 108 | 116 |
| Number of control observations | 276 | 276 | 276 | 276 | 276 | 276 | 276 | 276 |
| Standardized bias (p-value) | 0.919 | 0.796 | 0.974 | 0.935 | 0.961 | 0.919 | 0.958 | - |
| Treatment variable: ER | | | | | | | | |
| **[3] ATT** | **0.729**$^{**}$ | **0.594**$^{***}$ | **0.519**$^{**}$ | **0.577**$^{***}$ | **0.640**$^{***}$ | **0.613**$^{***}$ | **0.620**$^{***}$ | **0.305)**$^{***}$ |
| | (0.289) | (0.230) | (0.241) | (0.222) | (0.210) | (0.247) | (0.226) | (0.119) |
| Number of treated observations | 121 | 121 | 117 | 120 | 121 | 121 | 121 | 122 |
| Number of control observations | 270 | 270 | 270 | 270 | 270 | 270 | 270 | 270 |
| Standardized bias (p-value) | 0.898 | 0.977 | 0.953 | 0.954 | 0.989 | 0.898 | 0.987 | - |
| Treatment variable: DR | | | | | | | | |
| **[4] ATT** | **0.136** | **0.153** | **0.157** | **0.310** | **0.148** | **0.145** | **0.136** | **0.098** |
| | (0.384) | (0.383) | (0.360) | (0.100) | (0.452) | (0.346) | (0.397) | (0.134) |
| Number of treated observations | 90 | 90 | 88 | 90 | 90 | 90 | 90 | 90 |
| Number of control observations | 302 | 302 | 302 | 302 | 302 | 302 | 302 | 302 |
| Standardized bias (p-value) | 0.914 | 0.936 | 0.845 | 0.923 | 0.931 | 0.914 | 0.939 | - |

Note: country based cluster-bootstrapped standard errors in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. IPWRA stands for the Inverse-Probability-Weighted Regression Adjustment and uses saturated propensity scores.

**Appendix 10: an overview of GFPI construction steps**




[1] Considered with a negative sign.

Source: Authors

## Appendix 11: Identification of Structural Changes in Global Fiscal Performance

| *Country and period length filtered* | *Year(s) of Structural change* | *Confidence interval(s) (5% significance level)* |
|---|---|---|
| Austria | 2007 | [2006-2009] |
| Belgium | 2002; 2004; 2007; 2010 | [2001-2006]; [2003-2005]; [2006-2008]; [2009-2011] |
| Bulgaria | 2002; 2004 | [2000-2007]; [2003-2006] |
| Croatia | 2001; 2003; 2005; 2011 | [2000-2002]; [2002-2004]; [2003-2006]; [2000-2012] |
| Cyprus | 2008 | [2007-2010] |
| Czech Republic | 2001; 2003 | [2000-2002]; [2002-2013] |
| Denmark | 2003; 2007; 2010 | [2002-2004]; [2006-2009]; [2009-2012] |
| Estonia | 2008; 2011 | [2003-2010]; [2010-2013] |
| Finland | 2002; 2006; 2008 | [2001-2006]; [2005-2007]; [2007-2009] |
| France | 2008; 2010 | [2006-2011]; [2009-2013] |
| Germany | 2004; 2007; 2009 | [2003-2005]; [2006-2008]; [2007-2010] |
| Greece | 2010 | [2009-2013] |
| Hungary | 2006 | [2000-2008] |
| Ireland | 2007; 2010 | [2006-2008]; [2009-2012] |
| Italy | 2011 | [2005-2012] |
| Latvia | 2002; 2004 | [2000-2003]; [2002-2007] |
| Lithuania | 2001; 2008; 2011 | [2000-2003]; [2006-2009]; [2010-2013] |
| Luxembourg | 2004 | [2003-2005] |
| Malta | 2001 | [2000-2002] |
| Netherlands | 2001; 2003; 2008 | [2000-2002]; [2002-2004]; [2007-2009] |
| Poland | 2002; 2004 | [2001-2005]; [2000-2005] |
| Portugal | 2005; 2010 | [2004-2006]; [2004-2011] |
| Romania | 2001; 2004 | [2000-2002]; [2000-2005] |
| Slovak Republic | 2002; 2007 | [2000-2004]; [2000-2009] |
| Slovenia | 2008 | [2005-2013] |
| Spain | 2007; 2009; 2011 | [2006-2008]; [2008-2010]; [2010-2012] |
| Sweden | 2004; 2006 | [2000-2005]; [2005-2013] |
| United Kingdom | 2002 | [2000-2004] |

## Appendix 12: Rosenbaum's Sensitivity Analysis

| Treatment variable: FR | Nearest-neighbor Matching | | Radius Matching | | | local linear Matching | kernel Matching |
|---|---|---|---|---|---|---|---|
| | N = 1 | N = 3 | $r = 0.01$ | $r = 0.025$ | $r = 0.05$ | | |
| **[1] ATT-CAPB** | **0.698***[*] | **0.451***[*] | **0.676***[***] | **0.540***[**] | **0.515***[**] | **0.517***[***] | **0.510***[**] |
| | (0.493) | (0.375) | (0.339) | (0.342) | (0.263) | (0.288) | (0.326) |
| Γ | P-VALUE | P-VALUE | P-VALUE | P-VALUE | P-VALUE | P-VALUE | P-VALUE |
| 1.0 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 1.2 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 1.4 | 0,03 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 1.6 | 0,01 | 0,08 | 0,010 | 0,00 | 0,00 | 0,00 | 0,00 |
| 1.8 | 0,03 | 0,02 | 0,050 | 0,02 | 0,01 | 0,00 | 0,02 |
| 2.0 | 0,06 | 0,04 | 0,014 | 0,08 | 0,05 | 0,06 | 0,07 |
| **[2] ATT-GFPI** | **0.494***[***] | **0.537***[***] | **0.525***[***] | **0.504***[***] | **0.490***[***] | **0.499***[***] | **0.491***[***] |
| Γ | P-VALUE | P-VALUE | P-VALUE | P-VALUE | P-VALUE | P-VALUE | P-VALUE |
| 1.0 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 1.2 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 1.4 | 0,03 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 1.6 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 1.8 | 0,03 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 2.0 | 0,06 | 0,00 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 |

Note: Bootstrapped standard errors (with 500 replications) in brackets. *, **, *** indicate the significance level of 10%, 5%, and 1%, respectively. P-VALUE is the upper bond (sig+) of the Wilocoxon's signed rank test.

**Appendix 13: Correlations between the four indicators used to construct the GFPI based on Public Expenditure Growth Rate**

| | Public Expenditure Growth Rate | External Deficit | Variations of Fiscal Revenues | Sustainability Debt Index |
|---|---|---|---|---|
| Public Expenditure Growth Rate | 1.000 | 0.081 | 0.007 | 0.085 |
| External Deficit | - | 1.000 | -0.002 | 0.052 |
| Variations of Fiscal Revenues | - | - | 1.000 | 0.017 |
| Sustainability Debt Index | - | - | - | 1.000 |

**Appendix 14. LASSO : Least Absolute Shrinkage and Selection Operator**

The LASSO estimate is given by (see Tibshirani [1996])

$$\widehat{\beta}^{lasso} = \arg\min_{\beta} \left\{ \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{P} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{P} |\beta_j| \right\}, \quad (1.6)$$

where $\lambda > 0$ is a complexity (or tuning) parameter that controls the amount of shrinkage: the larger the value of $\lambda$, the greater the amount of shrinkage. The specific case where $\lambda = 0$ corresponds to the least squares fit. The value of $\lambda$ is found by grid search, and two criteria are generally retained to select the optimal value. The $\lambda_{min}$ approach considers the $\lambda$ value that yields the minimum mean cross-validated error, while the $\lambda_{se}$ approach considers the $\lambda$ value such as the error is within one standard error of the minimum error; in our analysis, we retained the latter approach.

# Bibliography

A. Afonso and S. Hauptmeier. Fiscal Behavior in the European Union. Rules, Fiscal Decentralization and Government Indebtedness. *ECB Working Paper Series No. 1054*, 2009. 31

A. Afonso, F. Huart, J. Tovar Jalles, and P. Stanek. Fiscal Rules and Twin Deficits: The Link between Fiscal and External Balances. *REM Working Paper 031-2018*, 2018. 34

I. Ahamada and P. Jolivaldt. Filtres usuels et filtre fondé sur ondelettes: étude comparative et application au cycle économique. *Economie et prévision*, page 149. 27

A. Alesina and G. Tabellini. A positive theory of fiscal deficits and government debt. *The Review of Economic Studies*, 57:403–414, 1990. 30

J. Alt, D. D. Lassen, and J. Wehner. It Isn't Just about Greece: Domestic Politics, Transparency and Fiscal Gimmickry in Europe. *British Journal of Political Science*, 44:707–716, 2014. 47

T.M. Andersen. Fiscal policy targeting under imperfect information. *Journal of International Money and Finance*, 34:114–130, 2013. 33, 42

H. Badinger. Fiscal rules, discretionary fiscal policy and macroeconomic stability: an empirical assessment for OECD countries. *Applied Economics*, 41:829–847, 2009. 26

H. Badinger and W. H. Reuter. The case for fiscal rules. *Economic Modelling*, 60:334–343, 2017. 52

H. Badinger, A. Fichet de Clairfontaine, and W. H. Reuter. Fiscal Rules and Twin Deficits: The Link between Fiscal and External Balances. *The World Economy*, 40:21–35, 2017. 34

M. Bamba, A. Minea, and J-L Combes. The effects of fiscal consolidations on the composition of government spending. *Applied Economics*, 52:1517–1532, 2020. 26, 30

M. Bergman, M.M. Hutchison, and S.E. Hougaard Jensen. Promoting sustainable public finances in the European Union: The role of fiscal rules and government efficiency. *European Journal of Political Economy*, 44:1–19, 2016. 27, 31, 50

U. M. Bergman and M. Hutchison. Economic stabilization in the post-crisis world:Are fiscal rules the answer? *Journal of International Money and Finance*, 52:82–101, 2015. 26

JM. Bertrand, E. Duflo, and S. Mullainathan. How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, 119:249–275, 2004. 46

O. Blanchard, A. Leandro, and J. Zettelmeyer. Redesigning the EU Fiscal Rules: From Rules to Standards. *Economic Policy*, 2021. 61

H. Bohn and R.P. Inman. Balanced Budget Rules and Public Deficits: Evidence from the U.S. States. NBER Working Paper No. 5533, 1996. 31

L. Bonatti and A. Cristini. Breaking the Stability Pact : Was it predictable ? *Journal of Policy Modelling*, 30:793–810, 2008. 40

M. Boschi and S. d'Addona. The stability of tax elasticities over the business cycle in European countries. *Fiscal Studies*, 40:175–210, 2019. 56

C. Calderon and K. Schmidt-Hebbel. The Choice of Fiscal Regimes in the World. *Central Bank of Chile Working Papers No. 487*, 2008. 40

M. Caliendo and S. Kopeinig. Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Survey*, 22, 2008. 41

M. Camdessus. International Financial and Monetary Stability. A Global Public Good. Chapter in Reforming the International Monetary and Financial System, P. B. Kenen and A. K. Swoboda, editors, IMF. 1999. 30

F. Caselli and J. Reynaud. Do Fiscal Rules Cause Better Fiscal Balances? A New Instrumental Variable Strategy. *IMF Working Paper No. WP/19/49*, 2019. 31

S. Cevik and K. Teksoz. Deep Roots of Fiscal Behavior. *IMF Working Paper No. 14/45*, 2014. 29

J-L. Combes, A. Minea, and M. Sow. Is fiscal policy always counter- (pro-) cyclical? The role of public debt and fiscal rules. *Economic Modelling*, 65:138–146, 2017. 26

J-L. Combes, X. Debrun, A. Minea, and R. Tapsoba. Inflation Targeting, Fiscal Rules and the Policy Mix: Cross-effects and Interactions. *The Economic Journal*, 128:2755–2784, 2018. 27, 31

European Commission. Report From The Commission To The European Parliament, The Council, The European Central Band And The European Economic And Social Committee: Alert Mechanism Report 2020. *COM/2019/651*, 2019. 33

Z. Darvas, P. Martin, and X. Ragot. European Fiscal Rules Require a Major Overhaul. *Notes du conseil d'analyse économique*, 47:1–12, 2018. 59

X. Debrun and L. Jonung. Under threat: Rules-based fiscal policy and how to preserve it. *European Journal of Political Economy*, 57:142–157, 2019. 62

X. Debrun and M.S. Kumar. The Discipline-Enhancing Role of Fiscal Institutions: Theory and Empirical Evidence. *IMF Working Paper No. 07/171*, 2007. 28, 31

X. Debrun, L. Moulin, A. Turrini, J. Ayuso i casals, and M. Kumar. Tied to the Mast? National Fiscal Rules in the European Union. *Economic Policy*, 23:297–362, 2008. 27, 28, 31, 50

K. M. Eklou and M. Joanis. Do Fiscal Rules Cause Fiscal Discipline over the Electoral Cycle? *IMF Working Paper No. WP/19/291*, 2019. 56

J. Escolano, L. Eyraud, J. Sarnes M. Moreno Badia, and A. Tuladhar. Fiscal Performance, Institutional Design and Decentralization in European Union Countries. *IMF Working Paper No. 12/45*, 2012. 28, 31

L. Eyraud, X. Debrun, A. Hodge, V. Duarte Lledo, and C. A Pattillo. Second-Generation Fiscal Rules : Balancing Simplicity, Flexibility, and Enforceability. *Staff Discussion Notes No. 18/04*, 2018. 62

A. Fatás and I. Mihov. The case for restricting fiscal policy discretion. *Quaterly Journal of Economics*, 118:1419–1447, 2003. 32

A. Fatás and I. Mihov. The macroeconomic effects of fiscal rules in the US states. *Journal of Public Economics*, 90:101–117, 2006. 32

A. Fedelino, A. Ivanova, and M. Horton. Computing Cyclically Adjusted Balances and Automatic Stabilizers. *Technical Notes and Manual*, 2009. 42

D. Foremny. Sub-national deficits in European Countries: The impact of fiscal rules and tax autonomy. *European Journal of Political Economy*, 34:86–110, 2014. 27, 31

N. Girouard and C. André. Measuring cyclically-adjusted budget balances for OECD countries. *OECD Economics Department Working Papers No. 434*, 2005. 42

B. Gootjes, J. de Haan, and R. Jong-A-Pin. Do Fiscal rules constrain political budget cycles? *DNB Working Paper No. 634*, 2019. 56

M. Guerguil, P. Mandon, and R. Tapsoba. Flexible fiscal rules and countercyclical fiscal policy. *Journal of Macroeconomics*, 52:189–220, 2017. 38, 40, 46

S. Guichard, M. Kennedy, E. Wurzel, and C. André. What Promotes Fiscal Consolidation: OECD Country Experiences. *OECD Economics Department Working Papers No. 553*, 2007. 26

M. Hallerberg, R. Strauch, and J. von Hagen. The design of fiscal rules and forms of governance in European Union countries . *European Jounral of Political Economy*, 23:338–359, 2007. 58

M. Hallerberg, R. Strauch, and J. von Hangen. Fiscal Governance in Europe. *Cambridge University Press*, 2009. 32

S. Hauptmeier and C. Kamps. Debt rule design in theory and practice - the SGP's debt benchmark revisited. *Presentation at ECFIN workshop "Fiscal rules in Europe"*, 2020. 59, 61

F. Heinemann, M-D. Moessinger, and M. Yeter. Do fiscal rules constrain fiscal policy? A meta-regression-analysis. *European Journal of Political Economy*, 51:69–92, 2018. 26, 27, 28, 32, 42

F. Holm-Hadulla, S. Hauptmeier, and P. Rother. The impact of expenditure rules on budgetary discipline over the cycle. *Applied Economics*, 44:3287–3296, 2012. 31

G.W. Imbens and J. Wooldridge. Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47:5–86, 2009. 46

O. Jorda and A. M. Taylor. The Time for Austerity: Estimating the Average Treatment Effect of Fiscal Policy. *The Economic Journal*, 126:219–255, 2016. 47

G. Kopits. Fiscal Rules: Useful Policy Framework or Unnecessary Ornament? *IMF working paper*, 2001. 25

G. Kopits and S.A. Symansky. Fiscal Policy Rules . *Occasional Paper No. 162*, 1998. 26, 28, 40

S. Krogstrup and S. Walti. Do fiscal rules cause budgetary outcomes? *Public Choice*, 136: 123–138, 2008. 46

S. Krogstrup and C. Wyplosz. A common pool theory of supranational deficit ceilings. . *European Economic Review*, 54:269–278, 2010. 30

M. S. Kumar and T. Ter-Minassian. Promoting Fiscal Discipline. *IMF Books*, 2007. 30

W. Marneffe, B. Van Aarle, W. Van Der Wielen, and L. Vereeck. The Impact of Fiscal Rules on Public Finances: Theory and Empirical Evidence for the Euro Area. *CESifo Working Paper No. 3303*, 2010. 28, 31

A. Minea and R. Tapsoba. Does Inflation targeting improve fiscal discipline? *Journal of International Money and Finance*, 40:185–203, 2014. 32

A. R. Mohanty and B. R. Mishra. Fiscal Performance Index of the States in India An Empirical Model Based Evidences. *Prajnan*, XLV, 2016. 33

S. Nickell. Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49:1417–1426, 1981. 33

M. Persson, T. Persson, and L. Svensson. Time consistency of fiscal and monetary policy: a solution. *Econometrica*, 74:193–212, 2006. 30

J. Poterba. Budget Institutions and fiscal Policy in the U.S. states. *American Economic Review*, 86:395–400, 1996. 46

Council Regulation. COUNCIL REGULATION (EC) No 1467/97 of 7 July 1997 on speeding up and clarifying the implementation of the excessive deficit procedure. *Official Journal of the European Union*. 40

W.H. Reuter. National numerical fiscal rules: not complied but still effective? *European Journal of Economy*, 39:67–81, 2015. 27, 28

P. R. Rosenbaum and D. B. Rubin. Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *The American Statistician*, 39(1):33–38, 1985. 46

P.R. Rosenbaum. Overt Bias in Observational Studies. *Observational Studies. Springer Series in Statistics.*, 2002. 46

P.R. Rosenbaum and D. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70:41–55, 1983. 40, 56

C. Sancak, J. Xing, and R. Velloso. Tax Revenue Response to the Business Cycle. *IMF Working Paper No. WP/10/71*, 2010. 56

R. Tapsoba. Do National Numerical Fiscal Rules really shape fiscal behaviours in developing countries? A treatment effect evaluation. *Economic Modelling*, 29:1356–1369, 2012. 27, 29, 31, 32, 40, 43, 47

R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58:267–288, 1996. 37, 76

D. Turner. Should Measures of Fiscal Stance be Adjusted for Terms of Trade Effects? *OECD Economics Department Working Papers No. 519*, 2006. 33

A. Velasco. Debts and deficits with fragmented fiscal policymaking. *Journal of Public Economics*, 76:105–125, 2000. 30

M. Villafuerte and P. Lopez-Murphy. Fiscal Policy in Oil Producing Countries During the Recent Oil Price Cycle. *IMF Working Paper No. WP/10/28*, 2010. 33

R. Wiese, R. Jong-A-Pin, and J. de Haan. Can successful fiscal adjustments only be achieved by spending cuts? *Economic Policy*, 54:145–166, 2018. 56

C. Wyplosz. Fiscal Rules: Theoretical Issues and Historical Experiences. Chapter in NBER book Fiscal Policy after the Financial Crisis (2013), Alberto Alesina and Francesco Giavazzi, editors. *Chicago: University of Chicago Press*, 85:495–525, 2013. 30

# Chapter 2

# Forecasting the Stability and Growth Pact compliance: a risk-management framework for the EU supranational fiscal rule

"The widespread introduction of fiscal councils has been viewed as compliance enhancing, improving transparency and raising the likelihood that breaches of rules carry a reputational and political cost. Further improvements will require political incentives to be better aligned with rule compliance." Caselli et al. [2018]

*A version of this chapter, co-authored with A. Barbier-Gauchard (University of Strasbourg, France) and T. Papadimitriou (Univerity of Thrace, Greece), is currently submitted in World Economy.*

## 2.1 Introduction

According to previous discussion in Chapter 1, enhancing fiscal discipline in the eurozone has become the bone of contention between the European authorities and the EU member states. Consequently, fiscal discipline is achieved, in the European context, by the Stability and Growth Pact (SGP). The first version of the SGP (1996) summarized the public finance criteria (in particular the criteria on public deficit which should be below 3% of GDP) introduced by the Maastricht Treaty (1992) and used them as rules to comply with by member states.

Despite this supranational fiscal rule, the eurozone has experienced several economic crises such as the first crisis in 2004, the Great Recession from 2007 to 2009 followed by the sovereign debt crisis and then the COVID 19 pandemic crisis. These events have systematically challenged the fiscal discipline. Consequently, the SGP was considered imperfect and thus reformed. Following the reforms of 2005, 2011 (Six-Pack) and 2013 (Two-Pack), the SGP has turned into a catalog of indicators that member states are expected to monitor without failures to ensure real coercive disciplinary power over the member states. As first intention, this set of fiscal rules aims at two complementary objectives: the "stability" of public finance on the one hand, requesting sound management of public finance and the "economic growth" in the EMU on the other, ensuring that national governments have enough leeway to intervene when deemed necessary (especially if a cyclical shock occurs). The SGP tries to achieve these targets using two instruments, namely the "dissuasive" arm intended to ensure strict compliance with the rule[1] and the "preventive" arm designed to encourage member states to present balanced and sound public finance in the medium term[2]. However, the increasing complexity in the SGP thus makes a link with the "fiscal rules trilemma" Debrun and Jonung [2019], discussed in the general introduction of this thesis. Indeed, to better understand the strengths and weaknesses of fiscal discipline in the Euro area, it is essential to back to the work of Kopits and Symansky [1998] and the definition an ideal fiscal rule which should be credible to conduct enforceability (it thus refers to the compliance[3] with the fiscal rule). Nevertheless, the Kopits and

[1] The disuasive arm consists in public deficit ceiling with sanctions imposed in the case of non-compliance, and exceptions to the rule in very specific economic circumstances.

[2] The preventive arm corresponds to a multilateral surveillance procedure with "stability programs", multi-annual programs setting fiscal guidelines over 3 years and making it possible to have visibility on public finance for the next 3 years in order to reach budget balance in the medium term.

[3] Fiscal rule compliance could be defined as the ability of the relevant fiscal aggregates (the budget bal-

Symansky [1998] criteria also referred to the "simplicity" or clarity of the fiscal rule which is not the case for the SGP. Finally, one key question is how to achieve compliance with such complex SGP?

The purpose of Chapter 2 is not to question the justification of the SGP and fiscal rules in general, nor to assess the fiscal rules effectiveness[4]. All these considerations were already adressed by Chapter 1. This second chapter also does not propose new fiscal rules to replace SGP[5] but it tries to fill a gap in the fiscal rules compliance literature by proposing a risk-management framework to strengthen the preventive arm of the Stability and Growth Pact (1996) in the Euro area. In that sense, Chapter 2 addresses the following question: "How to forecast fiscal rules compliance to strengthen fiscal rules performance monitoring?". To do so, this chapter focused on the forecast of the compliance with the SGP for all the 28 EU members between 2006 and 2018. It used a two-step Machine Learning approach that first consists in retrieving the key determinants of SGP compliance and then building a model to forecast the compliance of SGP, based on these major features.

This chapter thus enters in the literature on fiscal rules compliance. Indeed, on the one hand, a wide field of fiscal rules compliance literature focuses on the compliance score and the main determinants of fiscal rules compliance. Delgado-Téllez et al. [2017] for Spain regions using first-difference General Method of Moments or Reuter [2019] for EU member states and Nandelenga and Ellyne [2020] for sub-Saharan African countries, both used a logistic model and the highlighted determinants are mainly rule-related[6]. Larch and Santacroce [2020] provide highlights on correlations between the fiscal rules included in the SGP and various macroeconomic variables such as the market volatility index, the output gap, the nominal GDP growth or the quality of governance. The studies by Reuter [2019] and Nandelenga and Ellyne [2020] are seeking for causality, while Larch and Santacroce [2020] propose a simple correlation analysis. All these approaches consid-

ance, the debt-to-GDP ratio or government expenditure) to reach, in purely quantitative terms, the target set by the fiscal rules. In other words, compliance assesses whether or not the rule has been complied with and, in most cases, this assessment does not take into account escape clauses.

[4]See, for instance, Foremny [2014], Sacchi and Salotti [2015], Bergman et al. [2016], or also Barbier-Gauchard et al. [2021].

[5]As such, see papers on the second generation of fiscal rules as underlined by Eyraud et al. [2018] and Caselli et al. [2018] and also Darvas et al. [2018], Hauptmeier and Kamps [2020] or Debrun and Jonung [2019].

[6]The features that could strenghten fiscal rules compliance: registration in the law, level of rigor, degree of public finance coverage, etc.

ered contemporaneous information in the variable set and cannot be used for forecasting. Moreover, Reuter [2019] and Nandelenga and Ellyne [2020] created a universal model to investigate the compliance with a varying rule: in some countries the rule is a limit on the structural balance, in others the limit is set in the overall balance or the balance excluding public investment etc[7]. In this chapter, we will focus on the "3 % limit on public deficit" which is applied to all the European Union member states.

On the other hand, the latest reform of the Stability and Growth Pact undertaken with the Six Pack (2011) modified in depth the preventive arm by introducing the Macroeconomic Imbalance Scoreboard (MIP). The purpose of this scoreboard is to monitor a wide range of indicators used to identify any risk of internal and external imbalances that could destabilize public finance. The implementation of the European Semester, since 2010, was considered, at the time, a milestone towards efficient monitoring of public finance. While it is true that the SGP reforms introduced some powerful tools for close and thorough monitoring of public finance, these tools created new constraint for the countries : an excessively complex framework to fulfill. Instead of complicating the SGP, an alternative solution would be to simplify the rules in the monitoring process[8]. Improving the monitoring procedure could be interpreted as making the preventive arm more efficient, while simultaneously strengthening the dissuasive arm. In other words, strengthening the preventive arm may reduce the number of excessive deficit situations and thus strengthen the dissuasive arm. Following this idea, in this chapter, we propose a model to forecast the compliance with the 3% limit of public deficit. The model is created using Machine Learning, a methodological path rather unexplored in Macroeconomics, that often outperforms traditional Econometrics (see Ince and Trafalis [2006], Plakandaras et al. [2013]).

This chapter offers an original contribution in several ways:

First, this chapter focuses on the compliance with the European supranational fiscal rules introduced with the Stability and Growth Pact (1996) using Machine Learning. The Macroeconomic studies that use methods from the Machine Learning arsenal are still few (with a very positive trend though). The coupling of the SGP compliance forecasting with Machine Learning (ML) is novel and original. Second, the proposed method-

[7] For example, Denmark and the United Kingdom set a budget balance rule in 2005. However, the UK introduced a Golden rule, whereas Denmark set a general 2% GDP threshold on general government surpluses.

[8] The simplification of the fiscal framework is advocated by the European Commission or the European Fiscal Board [2020].

ology can handle the inherently limited in size datasets that Macroeconomics produce. Initially, Machine Learning systems required extensive datasets that were unavailable in Economics (with the exception of Finance). Today, the use of many new ML architectures that do not require unreasonably long data sets (like Support Vector Machines, Random Forests, and all the boosting-based methods), is an interesting and very promising avenue in Economic forecasting. Third, this contribution creates a very promising by-product: the geometric representation of the input set with the separation line. We do not just accurately forecast the compliance with the 3% of a EU country, but we also measure its distance from the separation line. The distance can be used a) to estimate the confidence of our forecast (the farthest from the line, the more confident we are), and b) to measure the recommendations and policies that eventually will change the forecasted negative fate of a country.

Our analysis focuses on the 28 EU member states over the period 2006-2018 and offers a new perspective in the debate of the dissuasive vs preventive arm: we propose a less complex but powerful preventive arm, that will minimize the cases that the dissuasive arm is needed. Indeed, this methodology can be used to design the delivered recommendations from the central authority to the "non-complier" countries. Instead of sanctions, the central authority may promptly (one year before) propose a set of well-targeted interventions that will change the fate of the "non-complier" (interventions that will change its position in the feature space: from the non-compliance subspace, to the compliance subspace).

The rest of the chapter is structured as follows. Section 2 presents the literature review. Section 3 exposes the data and the descriptive statistics. Section 4 describes the empirical strategy and the robustness approaches. Finally Section 5 reports the results and Section 6 concludes the chapter.

## 2.2 Literature review on fiscal rules compliance

The literature on fiscal rules compliance assessment dates back to the work of Reuter [2015]. However, it is closely linked to an older literature initiated by the seminal work of Kopits and Symansky [1998], which dealed with the qualities that fiscal rules should have. Very quickly, many authors stressed that is was impossible to define a fiscal rule that satisfies all these criteria simultaneously. In particular, Debrun and Jonung [2019] highlighted the "fiscal rules trilemma". They shown that with the current fiscal rules it is impossible to reconcile simultaneously three of Kopits and Symansky's criteria: (i) simplicity, (ii) flexibility, (iii) compliance.

A widespread literature focuses on the performance assessment of fiscal rules. Indeed, fiscal rules constitute a major tool to control fiscal discipline[9]. The starting point for fiscal rules assessment comes from Kopits and Symansky [1998]'s "ideal fiscal rule"[10]. Since fiscal rules are heterogenous through their design and application, they also present heterogenous effect and compliance. Following Kopits and Symansky [1998], some papers proposed a ranking of fiscal policy rules based on these ideal properties[11], and many others such Debrun et al. [2008] used empirical strategy and showed that national fiscal rules seem correlated to government fiscal performance (see also IMF [2009]). This theory is supported by other empirical studies, like Reuter [2015] using Least Square Dummy Variable, Bergman et al. [2016] with the system-GMM, or Tapsoba [2012] and Barbier-Gauchard et al. [2021] with the Propensity-Score Matching method. Similar conclusions were found on subnational level in Foremny [2014]. Fiscal rules performance is also relative to their macro-stabilizing power. For instance, Sacchi and Salotti [2015] highlighted that national fiscal rules contributed to the GDP stabilization. Guerguil et al. [2017] showed that flexible budget balance rules supported public expenditure stabilization (for standard definition of flexible rules see Schick [2010], Dabán [2011] or Caselli et al. [2018]). Numerous papers studied the impact of the supranational fiscal rule of the SGP on the counter-cyclical feature of national fiscal policy, as recently shown by Larch et al. [2021].

Another field of research for the fiscal rules investigates ways to resolve the "fiscal rules trilemma". This literature has inspired to the second generation of fiscal rules as underlined by Eyraud et al. [2018] and Caselli et al. [2018] which promote rule-based on fiscal frameworks and stronger incentives to reach compliance. The "fiscal Taylor rule" proposal by Debrun and Jonung [2019] offers an illustration of what a second generation of fiscal rule could be. In the same vein, Blanchard et al. [2020] propose to shift from fiscal

[9]Fiscal discipline is a wide concept including the whole fiscal framework. It should promote sound management of public finance. Fiscal discipline could be established through various fiscal rules.

[10]Indeed, Kopits and Symansky [1998] defined the "ideal fiscal rule" that must satisfy all these properties: (1) Suitability for the intended objective, (2) Clear definition, (3) General consistency, (4) Robust analytical foundations, (5) Transparency, (6) Simplicity, (7) Flexibility, (8) Credibility.

[11]See Creel [2003] for an attempt to evaluate these properties.

rules to "enforceable fiscal standards."

At the same time, other studies focused on existing fiscal rules and investigated the enforceability criteria (also appearing in the "fiscal rules trilemma"), which ultimately influences the degree of credibility attributed to the fiscal rules and refers to the compliance with the fiscal rule. This chapter takes its place in this field of literature, which implicitly assumes that: (i) the existence of fiscal rules is justified (see for instance Debrun et al. [2008] or Wyplosz [2012] for a general overview of the main reasons to introduce fiscal rules), (ii) the applied numerical limits are optimal. It is not the purpose of this chapter to question any of these hypotheses.

The studies on fiscal rules compliance are numerous. Some of them try to assess the compliance with fiscal rules based on the numerical fiscal rules databases published by the European Commission [2017] and by the IMF (2017)[12]. These databases provide information in terms of description and definition of the fiscal rule and its coverage, its statutory base, monitoring bodies, correction mechanisms in case of deviation from the rule, as well as experience with the respect of the rule.

Composite indicators are defined according to this information to assess the potential coercive power of fiscal rules: the Fiscal Rule Index (FRI) proposed by the European Commission [13] or by the IMF. However, in order to assess the effective coercive power of the fiscal rule (ie the effective compliance), the effective level of relevant fiscal aggregates should be compared to the numerical limit of fiscal rules. Reuter [2015] or Larch and Santacroce [2020] show that numerical fiscal rules are respected in 50% of cases. In the same vein, Delgado-Téllez et al. [2017] analyse the compliance on the subnational level in Spain and Cordes et al. [2015] focus on public expenditure rules compliance in advanced and emerging countries.

Other studies analyse the key determinants of fiscal rules compliance. Reuter [2019] looked at the determinants of fiscal rules compliance in the European Union from 1995 to 2005. This study showed that the rule specific features (in particular its legal basis and the existence of independent monitoring and enforcement authorities), the degree of government fragmentation or the political cycle have a significant influence on whether or not the national fiscal rule is respected. Nandelenga and Ellyne [2020] implemented a similar

[12] Schaechter et al. [2012].

[13] The Fiscal Rule Index (FRI) of the European Commission is calculated taking into account five criteria : 1) legal base, 2) binding character, 3) bodies monitoring compliance, 4) correction mechanisms, and 5) resilience to shocks.

analysis for the sub-Saharan African countries. However, neither the economic environment of the country (output gap, inflation rate, public debt or interest payments) nor the position in the economic cycle seem to play a role in national fiscal rule compliance. Moreover, combinations of fiscal rules (at national level or in addition with fiscal rules at regional or local level) do not significantly affect the compliance. Larch and Santacroce [2020] study the determinants of compliance with the supranational fiscal rule that exists in the European Union, introduced since the Stability and Growth Pact (1996). The European fiscal rules, which have been reformed several times, present different targets in terms of fiscal aggregates (deficit rule, debt rule, structural balance and expenditure). Their study covers the European Union countries from 1998 to 2019 and brings to light stark and persistent differences across countries. Their results reveal noteworthy links between numerical compliance in the one hand and some key macroeconomic variables (especially episodes of pro-cyclical fiscal policy) and the quality of governing institutions on the other (countries with “watchdogs”(Debrun et al. [2019]), i.e. national independent fiscal institutions). Nevertheless, as suggested by Reuter [2015], fiscal rules could be considered as a tool “to force governments to adjust their budgetary plans in such a way that the constrained variable is moving in the direction of the constraint”. In this case, the compliance degree and the factors explaining compliance may be considered as elements that make the forecasting of fiscal rules compliance possible.

The aim of this chapter is to offer an additional insight into the preventive arm of the fiscal rules in the eurozone. We propose to deepen the analysis in this direction by identifying the determinants of the SGP compliance and use them to forecast it. In this study, we propose a Machine Learning approach based on forecasting model targeting the supranational fiscal rule compliance. To the best of our knowledge, there is no other study in the literature investigating this topic under the Machine Learning perspective. Indeed, Machine Learning methodologies for classification and forecasting are increasingly applied in Economics problems.Gogas et al. [2015] were interested in the ability of the yield curve to forecast economic activity. They forecasted the positive and negative derivations of the real US GDP from its long-run trend over the period going from 1976 to 2014. Results showed that the best SVM model outperformed the econometric one (probit model). Gogas et al. [2018] used SVM in Forecasting U.S. Bank Failures and obtained a striking 99.22% overall forecasting accuracy, outperforming the well-established Ohlson’s score. Härdle et al. [2009] studied the default risk of companies with SVM and

Huang et al. [2004] used SVM in forecasting corporate credit ratings for the U.S. and Taiwan. They compared SVM to back propagation neural networks (BPNN); in every case the linear SVM outperformed the competition. This chapter is interested in extending the application range of Machine Learning to Public Policy Issues.

## 2.3 Data and descriptive statistics on the EU supranational fiscal rule and its determinants

### 2.3.1 Compliance with the Supranational Fiscal Rule in the European Union and Public Finance Statistics

In the European context, the concept of supranational fiscal rule appears in the Maastricht Treaty (1992) which launched the project of the creation of the Monetary Union and set the conditions to be satisfied to achieve it. Some of these relate to the stability of public finance that any candidate country should achieve to be accepted in the eurozone: a) the public deficit should not exceed the 3% of the GDP and b) the public debt should not exceed the 60% of GDP. As soon as a candidate country is admitted to the eurozone, it must satisfy the rules of the Stability and Growth Pact (1996) which initially only related to the 3% deficit threshold[14]. In the early 2000s, despite the supranational rules, some countries do not respect the fiscal rule. Excessive deficit procedures were launched against Portugal (in 2002), France and Germany (in 2003) but sanctions were never applied despite non compliance with the rule.

Several reforms (the reform of 2005, the Six Pack in 2011, the Two Pack in 2013) subsequently attempted to strengthen both the preventive arm (be able to have public finance

[14]Nevertheless, EU-members (both eurozone and non-eurozone members) are concerned of the SGP compliance since the European Commission requests multiannual programs on public finance. These programs provide a forecast on the level and nature of the public finance for the next 3 years. EU countries that don't belong to the eurozone, are expected to provide "stability programs" every year; eurozone countries must provide "convergence programs". Public finance of all EU countries are thus monitored. In the event of bad public finance trajectories, the European Commission will provide recommendations so that the States rectify the deficiencies. No deadline was initially imposed for these programs and monitoring was not as thorough as in the SGP's latest version.

at balance in the medium term) and the dissuasive arm (rules to respect and sanctions[15]) of the Stability and Growth Pact. The main idea is to foster public budget balance in the medium term to ensures the compliance with the rule of 3% in the event of deterioration of the economic situation. Thus, the monitoring of national cyclically adjusted public balance and cyclical features has been reinforced in recent years. One of the key objective of the Macroeconomic Imbalances Procedure (MIP) scoreboard introduced with the Six Pack (2011) is to identify any risk of internal and external imbalances in the country that could destabilize public finance for a long time. This scoreboard covers primary and auxiliary indicators: i) external imbalances and competitiveness indicators: current account balance (3 year average), net external investment position (in % of GDP), real effective exchange rate (3 year % change), export market shares (5 year % change) and nominal unit labor cost (3 year % change); ii) internal imbalances indicators: house price index deflated (1 year % change), private sector credit flow consolidated (% of GDP), private sector debt consolidated (% of GDP), general government sector debt (% of GDP), unemployment rate (3 year average), total financial sector liabilities non-consolidated (1 year % change).

Unfortunately all these measures seem insufficient to assess the risk for a country of exceeding the 3% threshold: many countries have continued to violate the rule as shown in Figure 1 and Table 1. We are interested in the preventive instrument, searching for the best indicators to forecast the 3% rule compliance. We are thus concerned on the reasons the SGP is still not satisfied after all its reforms. In this chapter, we focus our analysis on the 28 EU members over the period 2006-2018 which follows the first SGP reform and includes the Six-Pack (2011) and the Two-Pack (2013) reforms. Such choice allows us to inspect the government efforts in response to the SGP's reforms[16].

Figure 1 plots the SGP compliance of the 28 European countries between 2006 and 2018. It highlights high heterogeneities in government behavior regarding the SGP. As pointed out by the European Commission[17], the European Fiscal Framework and the SGP have become too complex. It appears difficult for a country to comply with all the SGP rules at the same time. The task of implementing a fiscal policy that takes care of all the MIP indicators and complies with the SGP goals simultaneously seems hardly possible.

[15] ranging between 0.5% and 2% of the GDP if noncompliance is observed

[16] Many macroeconomic variables that we use are not complete until after 2005, and our algorithm is very sensitive to missing values.

[17] See the European Commission website and communication on EU governance review

To make it worse, since the early 90's, national fiscal rules in the EU have substantially increased, adding one more layer of rules to comply with.


Note: "0" means SGP non-compliance and "1" means SGP compliance.

Figure 2.1: **SGP compliance in the 28 EU countries between 2006 and 2018**

Table 2.1: **Public Finance Statistics in European Countries between 2004 and 2018**

| Key indicator | **Mean** | **Country with best value** | **Country with worst value** |
|---|---|---|---|
| 3 % limit compliance (in % ) | 63,70 | Estonia, Luxembourg, Sweden (complied with the rule each year) | France (complied with the rule only 4 times) |
| Public Budget Balance (in % GDP) | -2,556 | Finland in 2008: 5,129 (highest public balance over the period) | Ireland in 2011: -32,028 (highest public deficit over the period) |
| General government gross public debt (in % GDP) | 58,69 | Estonia in 2009: 3,664 (lowest public debt over the period) | Greece in 2017: 183,45 (highest public debt over the period) |
| Gross fixed capital formation (in % GDP) | 21,96 | Slovak Republic in 2009: 37,4 (highest GFCF over the period) | Hungary in 2015: 11,5 (lowest GFCF over the period) |

Table 2.1 depicts an overview of public finance statistics and the SGP compliance. This heterogeneity of public finance reinforces what we found in Figure 1: it appears difficult for every member states to behave identically towards the SGP. Countries as Luxembourg, Estonia or Sweden complied with the SGP during the period under study while France succeeded only 4 times (in a third of the cases). As the member states react differently to the same symmetrical shocks (Frenkel and Nickel [2005], Velickovski and Stojkov [2014], Bk and Maciejewski [2017]), they also react differently to a single and general fiscal rule. The European Commission is already applying the idea during the European Semester to provide country-specific recommandations for public finance plans.

Even if the SGP failed in 36% of the cases[18], the sanctions were never applied in order to avoid the worsening of the economic situation of the Member State under scrutiny. Ireland highlighted a 32% of GDP public deficit during the Sovereign Debt Crisis and financial sanctions were never applied in the event of such difficulties. The major problem, if we let the deficit slip away, is that the debt can become too large. For example, the public debt of Greece was close to 200% in 2017, putting the EMU under the risk of a domino effect.

We can derive two conclusions from these findings: i) the subpar performance of the SGP is a direct indication that the current form of the SGP monitoring should change, and ii) the tools of the dissuasive arm cannot be applied for fear of worsening the macroeconomic status of the EU member under control. A simple solution would be to improve the monitoring and revise the recommendations. We could improve the preventive arm of the SGP, focusing on the forecasting and monitoring process.

So, in our analysis we consider that i) a simpler rule would be easily maintained by the member states, ii) the focus should be placed on the preventive arm, iii) the key features that lead to the non-compliance should be identified. In these lines, we focus only on the initial and simpler "SGP 1.0" (as called in Debrun and Jonung [2019] which corresponds only to the 3% limit on public deficit) compliance. The proposed forecasting methodology has two steps : i) the identification of the key features for compliance using a feature selection procedure, ii) the training of a Machine Learning model that can accurately fore-

[18] This finding about the SGP compliance is not surprising since similar findings exist for fiscal rules compliance in national level. Both show poor compliance. Despite the "Magnet-effect " of national fiscal rules, Eyraud et al. [2018] pointed out the "poor track record of compliance" with fiscal rules. Similarly, Reuter [2015] showed that governments make efforts to move closer to their national fiscal rules limit but in the end, in just 51% of the cases they successfully comply with the fiscal rules.

cast the compliance with the rules one year in advance, giving the Member State enough time to change the outcome.

### 2.3.2 Fiscal rules compliance: potential predictors

We tried to create a dataset containing all the potential features for forecasting the SGP compliance. Table 2.2 describes all the variables in our dataset. We used data for the 28 European Union members for a period from 2006 to 2018. Our variables are divided into 3 groups namely Country Specific Variables, MIP scoreboard indicators, and other Macroeconomic Variables. Basically, the MIP scoreboard is a good starting point since it contains variables intended to prevent external and internal imbalances[19] and offers many complete series of macroeconomic variables. Nevertheless, the MIP scoreboard main objective is not to forecast the SGP compliance. We have thus supplemented our dataset with the macroeconomic variables.

We take into account the country characteristics using the following Country Specific Variables: a Dummy variable reflecting if the country was an Advanced[20] country and a Dummy variable reflecting if the country was an Emerging country ; a Dummy variable reflecting if the country was a Resource-rich country, a Dummy variable indicating if the country was an EU Member in, a Dummy variable reflecting if the country was a Federal Country (X1 to X5). With the Dummy variable reflecting if the country was a eurozone member (X6) we check if the eurozone members comply with the rule more often than the non-eurozone members. We also checked if the formal procedure provided by the SGP makes a difference (X7).

The MIP scoreboard primary and auxiliary indicators are included using variables X8 to X38 and the macroeconomic variables that are not monitored by the European Commission for internal imbalance are variables X39 to X47.

X39 is a binary dummy variable indicating the presence of an economic crisis. This is a simple but broad indicator that captures all potential changes in an economy. Then, we follow Wiese et al. [2018] who proposed a measure for governments fiscal volatility using the Bai-Perron structural break filter. We thus test for the presence of structural breaks identified by the Bai and Perron test in structural balance for each country (X46).

[19] See Eurostat website's definition of MIP scoreboard

[20] IMF uses several criteria to elaborate countries classification. Among these, the three main ones are: per capita income level, export diversification and degree of integration into the global financial system.

Furthermore we use variables for Oil Prices, bonds yield, foreign currency and long-term sovereign debt ratings[21]. The fiscal space is simply measured as the difference between the public debt level of a country and the European Union median one (e.g. Cheng and Pitterle [2018] for an overview of fiscal space definitions). We have also introduced an indicator of the macroeconomic cycle (output gap is measured by the production function approach). In Reuter [2019] the output gap did not appear as an important determinant for national fiscal rule compliance while Larch and Santacroce [2020] highlighted a significant correlation between the output gap and SGP fiscal rules compliance. But the SGP provides escape clauses. We thus expect crises and cyclical fluctuations to have an impact on the forecasting of fiscal rules compliance.

The European Commission and the IMF rate the fiscal rules rigor (the fiscal rules rigor reflects its theoretical coercitive power) by proposing the Fiscal Rules Indices. The European Commission's national Fiscal Rules Strength Index (FRSI) consider the main features of fiscal rules: legal basis (is the rule written as a law or is it a government commitment?), level of public finance coverage (does the rule applies to all public administrations or just central government?), enforcement procedure (does the rule imposes sanctions?), the presence of a monitoring institution (is there an independent fiscal "watch dog" in charge of fiscal rules good-conduct?), stabilization power (does the rule exclude public investment of cyclical components?). By applying a standardization procedure to these scores, the European Commission is able to provide the FRSI, a strength index for each national (and subnational) fiscal rule. In a nutshell, from a methodological point of view, the European Commission calculates the FRSI, which measures the strength index of each fiscal rule separately, whereas the FRI (used in our study) provides an aggregate version of the strength index of all fiscal rules at all levels of government in a given country. We thus include the European Commission's FRI to test the hypothesis that countries implementing -in parallel- strong national fiscal rules are more likely to comply with the 3% rule of SGP. Following Annett [2000] and using the Database of Political Institutions from WorldBank, in variable X47 we calculate a measure of government fragmentation which reflects the dispersion of parties within the Parliament[22]. We want to check if government

[21] index from 1 to 21 coming from "A cross-Country Database of Fiscal Space" of World Bank (2019)

[22] We used the Annett [2000] definition of society fractionalization applying to Government fractionalization : $Fractionalization = 1 - \sum_{i=1}^{M} \left(\frac{n_i}{N}\right)^2, \quad i = 1, \ldots, M$

with N the total number of seats in the country parliament, $n_i$ is the number of seats belonging to the i-th party. Government fractionalization is thus defined as the probability that two randomly chosen deputies fragmentation is related to the SGP compliance. Y is our binary dependent variable describing if a country complies with the SGP (3% limit) in year t. We are trying to forecast $Y_t$ using lagged values of the 47 variables in our dataset $X_{i,j}$, $i = 1,...,47$, $j = t-1, t-2, t-3$.

come from two different parties (that also corresponds to World bank definition of government fragmentation).

Table 2.2: **Variables Overview**

| Variables | Correspondance Variables | Source/Database | |
|---|---|---|---|
| Y | Dummy variable =1 if 3% limit was complied in t | Authors' calculations | |
| X1 | Dummy variable reflecting if the country was an advanced country in t-p | IMF Fiscal Rules Database | Country Specific Variables |
| X2 | Dummy variable reflecting if the country was an Emerging country in t-p | IMF Fiscal Rules Database | |
| X3 | Dummy variable reflecting if the country was a Ressource-rich country in t-p | IMF Fiscal Rules Database | |
| X4 | Dummy variable reflecting if the country was an EU membership in t-p | IMF Fiscal Rules Database | |
| X5 | Dummy variable reflecting if the country was a Federal Country in t-p | IMF Fiscal Rules Database | |
| X6 | Dummy variable for eurozone entrance in t-p | IMF Fiscal Rules Database | |
| X7 | Dummy variable reflecting if the country was submitted to an enforcement procedure related to the supranational fiscal rules in t-p | IMF Fiscal Rules Database | |
| X8 | Gross domestic product, deflator, in t-p | Eurostat | MIP Scoreboard Primary and Auxialiary indicators |
| X9 | Total investment in t-p | Eurostat | |
| X10 | Gross national savings in t-p | Eurostat | |
| X11 | Inflation, average consumer prices, in t-p | Eurostat | |
| X12 | Population in t-p | Eurostat | |
| X13 | General government revenue in t-p | Eurostat | |
| X14 | General government total expenditure in t-p | Eurostat | |
| X15 | General government net lending/borrowing in t-p | Eurostat | |
| X16 | General government gross debt in t-p | Eurostat | |
| X17 | Net External Positions in t-p | Eurostat | |
| X18 | Current account balance in t-p | Eurostat | |
| X19 | Current account balance variations over 3 years in t-p | Eurostat | |
| X20 | Real Effective Exchange Rate in t-p | Eurostat | |
| X21 | Global export market share -% change over 5 years - in t-p | Eurostat | |
| X22 | Nominal unit wage cost -% change over 3 years - in t-p | Eurostat | |
| X23 | Debt of private sector in t-p, consolidated -% of GDP | Eurostat | |
| X24 | Liabilities of the financial corporations sector, -% change over 1 year - in t-p | Eurostat | |
| X25 | Unemployment rate - 3-year average - in t-p | Eurostat | |
| X26 | Unemployment rate in t-p | Eurostat | |
| X27 | Gross domestic product (real GDP) -% change over 1 year - in t-p | Eurostat | |
| X28 | Gross fixed capital formation in t-p -% of GDP - | Eurostat | |
| X29 | Gross domestic expenditure on R & D in t-p -% of GDP - | Eurostat | |
| X30 | Direct investment in the reporting economy (flow) in t-p -% of GDP - | Eurostat | |
| X31 | Direct investment in the reporting economy (stocks) -% of GDP | Eurostat | |
| X32 | Net trade balance of energy products in t-p -% of GDP - | Eurostat | |
| X33 | Real effective exchange rate, Euro area trading partners -% change over 3 years | Eurostat | |
| X34 | Terms of trade (goods and services) -% change over 5 years - in t-p | Eurostat | |
| X35 | Market share of world exports, volumes -% change over 1 year - in t-p | Eurostat | |
| X36 | Labor productivity -% change over 1 year - in t-p | Eurostat | |
| X37 | Residential construction in t-p -% of GDP - | Eurostat | |
| X38 | Employment -% change over 1 year - in t-p | Eurostat | |
| X39 | Dummy variable reflecting if there is a Crisis in t-p | Author's research | Other Macroeconomic Variables |
| X40 | Output gap (production function approach) in t-p | AMECO Database | |
| X41 | Oil Prices in t-p | FED | |
| X42 | Bonds yield in t-p | | |
| X43 | Foreign currency long-term sovereign debt ratings, index from 1-21 , in t-p | *World Bank*[1] | |
| X44 | Fiscal Space in t-p | Author's calculations | |
| X45 | Fiscal Rules Index (by European Commission) in t-p | European Commission fiscal rules Database | |
| X46 | Structural Breaks in t-p | Author's calculations (using Bai and Perron test) | |
| X47 | Government fragmentation in t-p | *World Bank*[2] | |

Note: [1] A cross-Country Database of Fiscal Space, 2019.

[2] Database of Political Institutions.

Y is the Dependent variable. X are potential predictors tested in the feature selection step. All variables used as predictors are a p lagged variable. We test for $p = 1, 2, 3$ for each feature. 47 variables are included considering 3 lagged so 141 features are tested.

## 2.4 Empirical strategy: Feature selection and the Support Vector Machine (SVM)

We conducted three sets of tests. In each set we used the best SVM model and compared it with the best logit model using the same dataset. First, we used just the main primary indicators of the MIP scoreboard as input variables. We wanted to check if these indicators, which are able to prevent internal (and external) imbalances could also be related with the SGP compliance. Then we tested our framework with the complete dataset of Table 2.2. If the second model outperforms the first one, it will be a direct indication that the MIP scoreboard is not enough to forecast SGP compliance. In the third set of tests we couple a well established feature selection method with our forecasting scheme, to identify just the necessary variables for our model.

### 2.4.1 Forecasting algorithms

#### The logistic function

In our analysis, we conducted tests using both a traditional econometric method (Logit) and an emerging methodology in Economics from the Machine Learning field (Support Vector Machine). Our goal is to create the most accurate forecasting model.

Our forecasting problem is transformed into a binary classification setup: we must forecast whether the countries in our dataset will comply with the rule (class 1) or not (class 0) $Y \in \{0;1\}$. The goal of our system is to use the input variables to find the linear or non-linear separator that correctly classifies the cases.

The collected data are represented by $x_{i,j}, i = 1,\cdots,n$ and $j = 1,\cdots,m$, describing $n = 364$ datapoints with $m = 141$ features, arranged in vectors $\mathbf{x}_i = [x_{i,1},\cdots,x_{i,141}]^{\mathrm{T}}$. The logistic function constrains Y in a range of $(0,1)$ and uses the sigmoid function :

$$p(y_i = 1) = \pi_i = \frac{exp^{\hat{\mathbf{x}}_i^{\mathrm{T}}\boldsymbol{\beta}}}{1 + exp^{\hat{\mathbf{x}}_i^{\mathrm{T}}\boldsymbol{\beta}}} \tag{2.1}$$

where $\hat{\boldsymbol{x}}_i = [1,\mathbf{x}_i^{\mathrm{T}}]^{\mathrm{T}}$ corresponds to a feature-column (the 1 corresponds to the intercept of the regression) and $\boldsymbol{\beta}$ is the column vector of the regression coefficients.

The classifier produces a probability score between 0 and 1. When the probability is lower than 0.5 the datapoint is put into class 0; when the probability is higher or equal to 0.5 the datapoint is put into class 1.

The goal is to find the $\boldsymbol{\beta}$ according to $p(\mathrm{Y}|\mathrm{X})$ that most accurately classifies correctly the observed data points. The problem is equivalent to maximizing the product of the likelihood probabilities:

$$l(\boldsymbol{\beta}) = \sum_{i=1}^{n}[y_i log(\pi_i) + (1-y_i) log(1-\pi_i)] = \sum_{i=1}^{n}[\, y_i \; log(\frac{\pi_i}{1-\pi_i}) + log(1-\pi_i)] \qquad (2.2)$$

$$= \sum_{i=1}^{n}[y_i \hat{\boldsymbol{x}}_i^{\mathrm{T}} \boldsymbol{\beta} - log(1+exp^{\hat{\boldsymbol{x}}_i^{\mathrm{T}} \boldsymbol{\beta}})]$$

**The Support Vector Machine (SVM)**

**a) The Support Vector Machine in linearly separable cases**

SVM is a supervised Machine Learning method[23] for the binary classification of a set of data points. SVM aims at identifying a small subset of data points from the initial dataset, called Support Vectors, that define the position of the linear separator between the two classes.

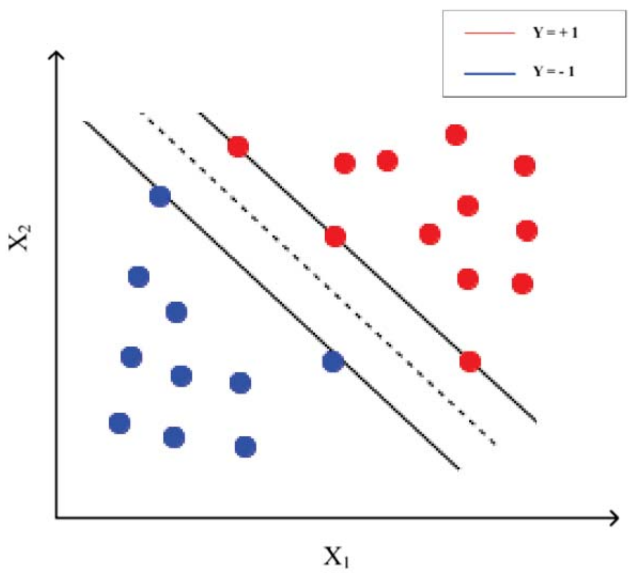



Figure 2.2: **Hyperplane and Support Vectors**

[23]Supervised learning is the concept where given a set of data and a set of observations, an algorithm creates a mapping function which describes the relationship from the data to the observations.

Consider $y_i$ as the binary outcome taking the value of -1 or 1 (in the logistic model $y_i$ takes the values 0 and 1). If the two classes are linearly separable, the separator is defined by:

$$f(\mathbf{x}_i) = \mathbf{w}^{\mathrm{T}}\mathbf{x}_i - b = 0 \tag{2.3}$$

where $\mathbf{x}_i$ is the $i-th$ $m$-sized data point (for our tests the datapoints are $i = 1, \cdots, 364$ and the features are $m = 141$) ; **w** is the weight vector, b is the bias. In that sense all data satisfy:

$$\mathbf{w}^{\mathrm{T}}\mathbf{x}_i - b > 0 \quad if \quad y_i = +1$$

$$\mathbf{w}^{\mathrm{T}}\mathbf{x}_i - b < 0 \quad if \quad y_i = -1, \qquad thus \quad y_i f(\mathbf{x}_i) > 0, \ \forall i$$

Ideally, the optimal separator is defined as the decision boundary that classifies each data point to the correct subspace and has the maximum distance from each class. This distance is often called "margin" and corresponds to the exact distance of the hyperplane with each class.

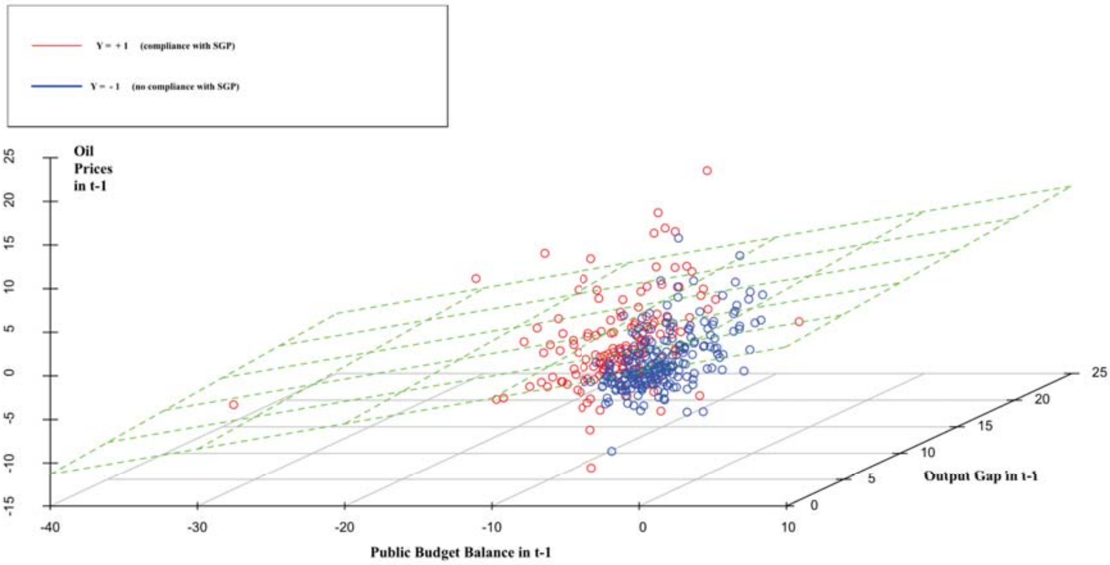



Figure 2.3: **Search for optimal linear separator hyperplane in the 3D data space of our dataset**

In Figure 2 and Figure 3 we provide a representation for the case of two and three dimensional systems[24]. The different colors of the data points correspond to the two classes of our dataset. In Figure 2 the linear separator corresponds to the dashed line, the margin

[24] In our case we have more than 3 variables/features so it is impossible to show the cloud of the datapoints in full. However the 3-d representation in Fig. 2.3 is created using three variables taken from our dataset.

lines corespond to the continuous lines and the Support Vectors correspond to the circles identified on the margin lines.

The separating hyperplane is identified using the Lagrange relaxation of a quadratic problem:

$$\min_{\mathbf{w},b}\max_{\mathbf{a}}\left(\frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{N} a_i[y_i(\mathbf{w}^{T}\mathbf{x}_i - b) - 1]\right) \tag{2.4}$$

In Equation (2.4) $\mathbf{a} = [a_1,..,a_n]^{T}$ correspond to the non-negative Lagrange multipliers. (2.4) is never used to calculate the solution. Instead we use the simpler dual problem described by:

$$\max_{\mathbf{a}}\left\{\sum_{i=1}^{N} a_i - \sum_{j=1}^{N}\sum_{k=1}^{N} a_j a_k y_j y_k \mathbf{x}_j^{T}\mathbf{x}_k\right\} \tag{2.5}$$

with $\sum_{i=1}^{N} a_i y_i = 0$ and $0 \leq a_i, \forall_i$. By solving (2.5) we obtain the location of the hyperplane given by:

$$\widehat{\mathbf{w}} = \sum_{i=1}^{N} a_i y_i \mathbf{x}_i \tag{2.6}$$

$$\widehat{b} = \widehat{\mathbf{w}}^{T}\mathbf{x}_i - y_i, i \in V, \tag{2.7}$$

where $V = \{i : 0 < a_i\}$ is the sert of support vector indices.

To consider a system contaminated by the presence of noise and outliers in the dataset Cortes and Vapnik [1995] introduced non-negative slack variables $\xi_i \geq 0, \forall i$ that can tolerate the misclassification of some cases. In order to keep the misclassification set as small as possible, each misclassification yields an additional financial cost in the objective function that we try to minimize.

$$\min_{wb\boldsymbol{\xi}}\max_{a\boldsymbol{\mu}}\left\{\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i - \sum_{j=1}^{N} a_j[y_j(\mathbf{w}^{T}\mathbf{x}_j - b) - 1 + \xi_j] - \sum_{k=1}^{N}\mu_k^{T}\xi_k\right\} \tag{2.8}$$

where the non-negative slack $\xi_i$ correspond to the distance of vector $\mathbf{x}_i$ from the hyperplane when classified erroneously. $\boldsymbol{\mu}_k = [\mu_1,...,\mu_n]$ are Lagrange multipliers. The optimal hyperplane is finally given by:

$$\widehat{\mathbf{w}} = \sum_{i=1}^{N} a_i y_i \mathbf{x}_i \tag{2.9}$$

$$\widehat{b} = \widehat{\mathbf{w}}^{\mathrm{T}} \mathbf{x}_i - y_i, i \in \mathrm{V}, \tag{2.10}$$

where $\mathrm{V} = \{i : 0 < a_i < \mathrm{C}\}$ is the set of support vector indices. Parameter C is found using power of 2 grid search and $2^{-7} \leq \mathrm{C} \leq 2^{7}$.

**b) The Support Vector Machine for the non linearly separable case**

Real world phenomena are often nonlinear. Linear models like the SVM are unable to model these systems correctly. To overcome the problem of nonlinearity the SVM paradigm is coupled with the kernel trick. Kernels project the initial data space to a feature space of higher dimensionality. Instead of searching for the optimal separator in the data space, we look for it in the feature space and return the solution to the initial data space (see Figure 4). So when the kernel is nonlinear and although the separator in the feature space is linear (SVM yields only linear separators), the separator in the data space is nonlinear. The kernel trick ensures low computational cost; the projection is performed in the inner product space [25], instead of projecting each point separately in the feature space. Introducing the kernel projection in the minimization of the objective function transforms it to:

$$\max_{\mathbf{a}} = \sum_{i=1}^{\mathrm{N}} a_i - \frac{1}{2} \sum_{j=1}^{\mathrm{N}} \sum_{k=1}^{\mathrm{N}} a_j a_k y_j y_k \mathrm{K}(\mathbf{x}_j, \mathbf{x}_k). \tag{2.11}$$

with $\sum_{i=1}^{\mathrm{N}} a_i yi = 0$ and $0 \leq a_i \leq \mathrm{C}$, $\forall_i$. In our tests we investigated the three most commonly used kernels:

$$Linear \qquad \mathrm{K}_1(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^{\mathrm{T}} \mathbf{x}_j + r, \tag{2.12}$$

$$\mathrm{RBF} \qquad \mathrm{K}_2(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2}, \tag{2.13}$$

$$Polynomial \qquad \mathrm{K}_3(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^{\mathrm{T}} \mathbf{x}_j + r)^d, \tag{2.14}$$

Now, the rule for classifying a data point **x** is given by:

[25] kernel functions are called “generalized dot products’

$$f(\mathbf{x}) = sign\left\{\sum_{i=1}^{\mathrm{N}} a_i y_i \mathrm{K}(\mathbf{x}_i, \mathbf{x}) + b\right\} \tag{2.15}$$

Indeed, if $f(\mathbf{x}) > 0$ the point is classified as belonging to class $+1$; otherwise, it is in class $-1$.

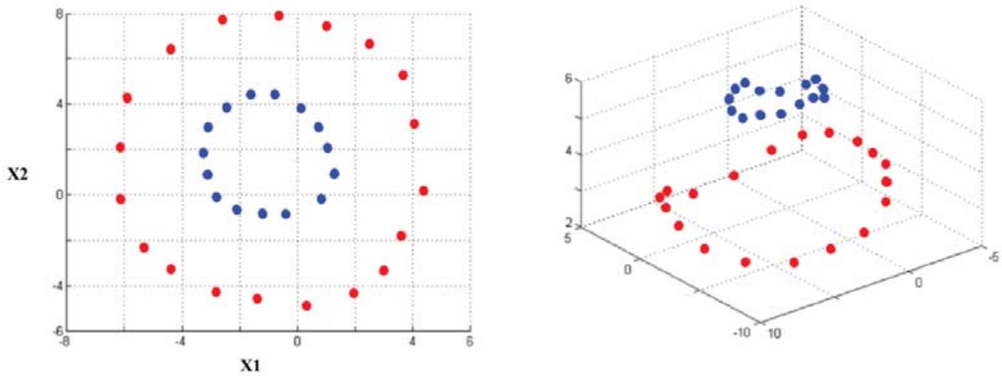



Figure 2.4: **Kernel projection to make the two classes linearly separable**

In Figure 4, the system in the left figure corresponds to a dataset of two non linearly separable classes. The system in the right is the projection of the same dataset in a 3D feature space that the two classes are linearly separable. For further details on SVM methodology see Gogas and Papadimitriou [2021].

### 2.4.2 Feature Selection: The logistic LASSO

The goal of feature selection is the reduction of input set, by removing irrelevant or redundant features for our model. By reducing the input set, we decrease the computational cost of training, and minimize the risk of model overfitting.

Friedman et al. [2009] proposed LASSO as a regularization alternative that overcomes the inability of ridge regression to reduce the number of predictors in the final model. LASSO applies a regularization process where the coefficients of some of the input variables are penalized and shrunk to zero. The main goal of the method is to minimize the prediction error, yielding as a by-product the feature selection of the variables.

The shrinkage operation identifies the key features from our dataset, avoiding the problem of transformation-based dimension reduction methodologies using Factor Anaysis, Principal Component Analysis or Independent Component Analysis, (to name but a few) which lead to factors that are uninterpretable.

Finally, the LASSO estimator applied in logistic regression is:

$$\hat{\beta}(\lambda) = \underset{\beta}{\operatorname{argmin}}(n^{-1}\sum_{i=1}^{n}\rho_{(\beta)}(X_i, Y_i) + \lambda||\beta||_1) \tag{2.16}$$

Parameter λ is found by grid search (view Appendix 1) and used the one-standard error rule. If there are more than one models with similar performance, we keep the most parsimonious one.

### 2.4.3 Measurement forecasting performance

The performance of our models is calculated using the forecasting accuracy defined as the ratio of the correctly forecasted observations over all the observations.

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{2.17}$$

where TP is the number of True Positive instances (correctly forecasted positive instances), TN is the number of True Negative instances (correctly forecasted negative instances), FP is the number of False Positive instances (incorrectly forecasted positive instances), FN is the number of False Negative instances (incorrectly forecasted negative instances). We remind that in our set-up a positive instance is a Member State that satisfied the rule in year t, while a negative instance describes the opposite case.

The forecasting accuracy is a simple and easy to use metric of the model's performance; nonetheless, it is a coarse and superficial measurement. Consider, for example, a dataset with 90 positive cases and 10 negative ones coupled with a naïve model yielding only positive forecasts. The accuracy of the model is 90%, which is quite misleading since it missed all the negative cases. The confusion matrix (Figure 5) is a deeper and richer representation of the model's performance, uncoupling the performance of the model in the two potential outcomes.

Indeed, a false positive case is damageable for the EU economy: if we incorrectly forecast that a country will comply with the SGP, no recommendations or measures will be prepared by the Commission since public finance are not expected to worsen. Too many false positive cases could jeopardize the sustainability of the entire currency area. So, it is important to create a forecasting model yielding the fewer possible false positives. This, however, should not produce the side effect of too many false negatives. The naïve example of the last paragraph describes such a trivial case. A false negative case, i.e., a country incorrectly forecasted to miss the rule, will force the Commission to recommend a set of unnecessary strict measures, that could harm the economy by reducing its fiscal capacity. So, it is important to assess the performance of the model in both the positive and the negative cases.

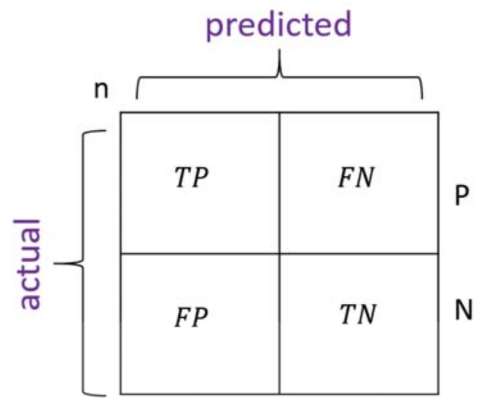



Figure 2.5: **The confusion matrix**

### 2.4.4 Robustness

Machine Learning methodologies are, in general, unaffected by the reverse causality bias, which is a common problem in classic econometrics. They suffer, though, from the curse of overfitting: a common error occurring when the model learns to describe the training data instead of the phenomenon at hand. Overfitting can be avoided using the hold-out validation and the more powerful K-fold cross validation approach. In hold-out validation the dataset is split into a 'training' set and 'test' set. The model is trained on the training set and the test set is used to evaluate the generalization performance of the model on unknown data. If the training accuracy is much higher than the testing accuracy, it is a strong indication that the model overfit the training dataset. Usually, we use around 80% of the data for the training and the rest for testing. K-fold cross validation repeats hold-out $k$ times. Indeed, our data set is split up into $k$ equally sized subsets and the training-testing steps are implemented $k$ times. At each turn, a different subset is used as the 'test' set, whereas the rest of the $k$-1 subsets are grouped and constitute the 'training' set. The average performance from every fold is used to obtain the optimal model.


Figure 2.6: **5-fold cross validation example**

## 2.5 Results

### 2.5.1 Forecasting the SGP compliance

We tested every dataset using the SVM classification setup coupled with three kernels (the linear, the RBF and Polynomial one). We used the performance of the logistic regression on the same datasets as a benchmark for our ML models. In the first step of our study, we trained our models using the primary indicators of the MIP scoreboard. The goal was to evaluate the ability of the MIP scoreboard to forecast the compliance with the 3% rule. Then we performed the same training scheme using the whole 141 variables dataset. The results in the first two sets are reported in Table 2.3.

Table 2.3: **"Compliance with 3% limit" forecasting accuracy: comparison of models (%)**

| Model | MIP scoreboard Primary indicators included | All features included (141) | MIP scoreboard Primary indicators included | All features included (141) |
|---|---|---|---|---|
| Linear SVM model | 64.6 | 83.5 | 85.1 | 87.0 |
| Quadratic SVM model | 69.5 | 83.0 | 79,6 | 88.9 |
| RBF SVM | 62.9 | 80.2 | 77,8 | 88.9 |
| Logistic model | 63.7 | 75.5 | 75.9 | 73.3 |
| validation method | k-Fold cross validation | K-fold cross validation | hold-out | hold-out |

Note: hold-out splits up dataset into a 'trainset' (85%) and 'testset' (15%). Results are on testset. k-Fold cross validation is a 5-Fold cross validation and gives mean results. Parameter C in SVM is equal to $2^1$ and obtained using power of 2 grid search.

Comparing the performance of the models trained on the two datasets, it is easy to verify that the full dataset models dominantly outperformed the MIP scoreboard primary indicators models in almost every case. In the MIP scoreboard dataset using the hold-out validation method the SVM coupled with the linear kernel achieved the top performance reaching 85.1% accuracy (the linear model fed with the full dataset using hold-out validation achieved 87%). In the full dataset the SVM models equipped with the non-linear kernels (the quadratic and the RBF kernel) using the hold-out validation both achieved 88.9% forecasting accuracy, which is the top performance achieved by any type of model using the hold-out validation. In the case of the cross-validation, the improvement of using the full dataset over the MIP scoreboard dataset is more impressive. The accuracy of the models using the MIP scoreboard on cross-validation ranges from 62.9% in the case of the RBF-SVM model, to 69.5% in the case of the Quadratic kernel-SVM model; the accuracy of the models using the full dataset on cross-validation ranges from 75.5% in the case of the logit model, to 83.5% in the case of the linear-SVM model (the top performance using the cross-validation). The hard evidence from the models' performance suggest that the full dataset has more forecasting power than the MIP scoreboard primary indicators. The next step of our study is to identify the optimal input set using the LASSO based feature selection method.

LASSO Feature selection[26] highlighted a set of 12 key variables that are essential to

[26] Appendix 1 reports the LASSO procedure results. Following the one-standard error rule in LASSO, 12 features were identified as important.

forecast SGP compliance. Following Gogas et al. [2018], these selected variables were introduced into the SVM forecasting model and we implement a shrinking procedure. We compared the set that includes the LASSO selected variables with all the sets generated by removing one variable from this set. We kept the optimal one and continued the procedure until no improvement could be achieved. A set of 8 features was identified from it (Table 2.4):

Table 2.4: **Best predictors:**

| |
|---|
| General government fiscal balance in t-1 |
| Liabilities of the financial corporations sector, % change over 1 year, in t-1 |
| Dummy variable reflecting if there was a crisis in t-1 and t-2 |
| output gap in t-1 |
| Oil prices in t-1 |
| Bond yield in t-1 |
| Fiscal space in t-1 |

The feature set is composed by a) the General Government Fiscal Balance in $t-1$ (this was to be expected since a degraded fiscal balance in one year, will eventually have an impact in the next one), b) liabilities of the financial corporations' sector in $t-1$ (the global financial crisis highlighted the dependence between the solvency of financial institutions, the quality of their liabilities and the public finance sustainability[27], c) the dummy variable reflecting the occurrence of a crisis in $t-1$ and $t-2$, indeed, economic crises have a double impact on public deficits: they induce economic recessions and they create increased investment needs (in addition, the identification of these two features is a direct indication that the SGP escape clause should be adapted to crises duration and not only focus on the fall of the GDP during recessions), d) output gap in $t-1$ (the output gap is an indicator of the position in the economic cycle – the increased GDP volatility in times of poor economic conditions impacts the public deficit and thus the SGP compliance), e) the oil price in $t-1$(the level of the oil price has led to crises directly, as in the case of 1973 and 1979, or by proxy, as in 2008 - the "yellow vests" movement, triggered in 2018 by the oil prices in France, revealed once more the consequences on the public deficit that such

[27]For one, we note that commercial banks hold large quantities of treasure bills

situations can create), f) bond yield in t-1 and g) fiscal space in $t-1$ (both variables are related to the fiscal flexibility of a government, especially in periods of crises - this is in line with the Romer and Romer [2018] study that highlighted the importance of fiscal space during financial crises and normal recessions). In Table 2.5, we report the performance of the feature set in every type of models and for both validation cases.

Table 2.5: **"Compliance with 3% limit" forecasting accuracy with only Best Predictors(%)**

| Model | Features selected by LASSO included | Features selected by LASSO included |
|---|---|---|
| Linear SVM model | 90.4 | 98.1 |
| Quadratic SVM model | 84.6 | 87.0 |
| RBF SVM ($\gamma = 12$) | 86.5 | 88.9 |
| Logistic model | 78.5 | 76.3 |
| validation method | k-Fold cross validation | hold-out |

Note: hold-out splits up dataset into a 'trainset' (85%) and 'testset' (15%). Results are on testset. k-Fold cross validation is a 5-Fold cross validation and gives mean results. Parameter C in SVM is equal to $2^1$ and obtained using power of 2 grid search.

The models created using the selected features achieved the top performance in both types of validation (98.1% in the case of Hold Out validation and 90.4% in the case of K-Fold cross validation) using the linear SVM model. We remind that the reported performance for cross validation is the mean testing accuracy of the 5 folds. Furthermore, we tested the top performing model in the whole dataset (all the observations) and achieved 91.7% forecasting accuracy. The confusion matrix of this case can be found in Figure 7.

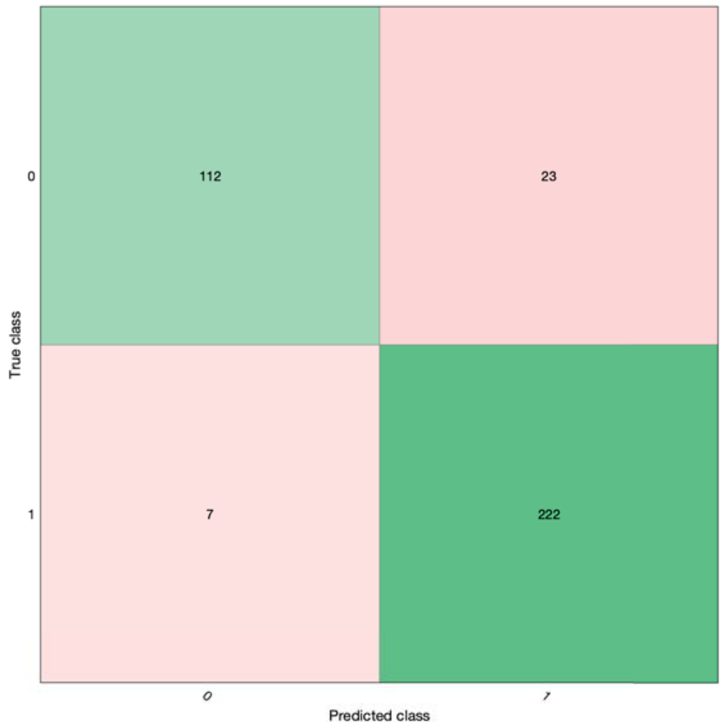



Figure 2.7: **Linear SVM confusion matrix (hold-out cross validation)**

The confusion matrix revealed that the model forecasted correctly 112 of the 135 negative cases that a Member State did not comply with the 3% rule, while it kept the false alarms in relatively low levels: 7 false alarms in 229 positive cases (a false alarm happens when the model inaccurately forecasts the non-compliance with the 3% rule). So, the model displays high accuracy in identifying the non-compliance, while keeping in low levels the cases that a Member State will be given recommendations for unfounded reasons (false alarms).

It may be risky to make assessments on the testing sets using the presented scenarios (the models may be slightly suboptimal due to the use of grid search on the identification of the model hyperparameters), but there is a strong indication that the main MIP Scoreboard indicators dataset was the least successful: it does not appear as appropriate for monitoring internal imbalances compared to the LASSO based feature set. Nevertheless, to be fair, the MIP Scoreboard indicators were not introduced, specifically, to monitor the SGP compliance.

Moreover, some of these indicators are defined over several years and increase the complexity of the monitoring system. In this sense, the identified features in our optimal input set could be used in a forecasting tool that will reinforce the European Fiscal framework surveillance.

If we try to analyse the model performance in a national level, we encounter the following missed negative cases[28]: Belgium (2009 and 2011), Bulgaria (2014), Croatia (2006, 2013 and 2015), Czech Republic (2012), Denmark (2012), Finland (2014), France (2016), Germany (2009 and 2010), Ireland (2008), Italy (2006 and 2011), Lithuania (2008), Malta (2012), Poland (2014), Slovak Republic (2006), Romania (2007 and 2018), Slovenia (2013), Spain (2008). The 12 out of the 23 cases involve the two big crises namely the Global Financial Crisis 2007-2009 and the Sovereign Debt Crisis 2010-2012. It must be noted that during these periods, the member states are more concerned in facing the direct implications of the crisis, than complying with the SGP rules. This change of macroeconomic aiming is usually unexpected and cannot be forecasted in the lagged instances of the feature variables. Let us remind that these crises led to the introduction of the escape clause in the Six-Pack in 2011.

### 2.5.2 Distance from separator hyperplane and policy implications

Even though the great majority of Machine Learning methodologies work in a "black box" framework (we never yield an analytical form of the forecasting model, as is usual in classic Econometrics), the proposed methodology is able to offer a classic analytical version of our model. Indeed, findings of Section 2.5.1 lead to the optimum forecasting model which corresponds to a linear decision boundary (also called separator hyperplane) that separates the SGP compliant observations from the SGP non-compliers with 91.7% accuracy using 8 explanatory variables. The analytical form of the separator hyperplane is:

$$\mathrm{H}: 3{,}546x_1 - 0{,}285x_2 - 0{,}338x_3 - 0{,}458x_4 - 1{,}014x_5 - 0{,}416x_6 - 0{,}581x_7 - 0{,}175x_8 + 0{,}793 = 0 \tag{2.18}$$

where $x_1$ corresponds to the General government fiscal balance in $t-1$, $x_2$ the Liabilities of the financial corporations sector ( % change over 1 year) in $t-1$, $x_3$ Crisis dummy in $t-1$, $x_4$ Crisis dummy in $t-2$, $x_5$ the output gap in $t-1$, $x_6$ the Oil prices in $t-1$, $x_7$ the Bond yield in $t-1$ and $x_8$ the Fiscal space in $t-1$ .

Through this identification we see several implications. First, we make the variables' impact interpretable. Indeed, the General government fiscal balance in $t-1$ is linked to a positive parameter. It is not surprising that an increase in the public balance affects posi-

[28]Missed positive cases are the following: Bulgaria (2009), Croatia (2008), Hungary (2012), Italy (2014), Malta (2013), Slovak Republic (2013), United Kingdom (2016).

tively the SGP compliance since the 3% rule is defined on the public budget balance. The parameter associated with the lagged value of Public Budget Balance is the largest one, reflecting that this is the strongest factor with the highest influence on SGP compliance. On the other hand, crisis dummy in $t-1$ and $t-2$ has a negative influence for an observation to be in the compliance subspace. 1-year change in Financial Sector Liabiliaties (in $t-1$) is associate with a negative parameter. It is a MIP scoreboard auxiliary indicator with an indicative threshold of 16.5%. Thus, if a Member State highlights this indicator above the limit, we are in the presence of a potential imbalance. We can therefore suggest that a significant increase in this indicator favors the SGP non-compliance. We already mentionned that the Oil prices could be interpreted as an advance indicator for economic crisis as the economic history suggests. It is therefore not surprising that this variable is negatively related with the SGP non-compliance. All deviations from the economic trend reflected by change in the output gap also destabilize the public finance and decrease the SGP compliance. Bond Yields are also associated with a negative parameter. Indeed Bond Yield increases with the debt sovereign risk default (in the EU for example, Greece has the highest sovereign risk premium), it therefore seems possible that countries with high Bond Yields to run higher deficit and comply with the SGP fever times. We also find a negative sign for the parameter relative to Fiscal Space suggesting that countries with more fiscal flexibility tend to increase public deficit and are expected to comply with the SGP more rarely.

Second, we can use the analytical form of equation 2.18 to calculate the distance between any point and the hyperplane. It is the distance that a country should be displaced to pass from the one subspace to the other. If a point is forecasted not to comply with the SGP, then the European Commission detailed recommendations should result in a displacement large enough to pass in other side. Obviously, between two "non-compliers" it is easier to change the "fate" of the country closer to the separation hyperplane, than of the one farther away. Similarly, we may use the distance in the case of a country forecasted to comply, to estimate a confidence parameter of the SGP compliance. A "complier" country close to the separator should be closely monitored, since a small perturbation in the economic system or a public budget failure may displace it in the non-compliance subspace. The same alertness is not needed in the case of a country forecasted to comply with the SGP with a large distance between its point and the separator hyperplane.

If we consider one observation A with coordinates $(x_A, y_A, z_A, r_A, s_A, t_A, v_A, w_A)$, its distance from the separator hyperplane, is defined as follow:

$$d(\mathrm{A}) = \frac{|\,3,546x_{\mathrm{A}} - 0,285y_{\mathrm{A}} - 0,338z_{\mathrm{A}} - 0,458r_{\mathrm{A}} - 1,014s_{\mathrm{A}} - 0,416t_{\mathrm{A}} - 0,581v_{\mathrm{A}} - 0,175w_{\mathrm{A}} + 0,793\,|}{\sqrt{3,546^2 + (-0,285)^2 + (-0,338)^2 + (-0,458)^2 + (-1,014)^2 + (-0,416)^2 + (-0,581)^2 + (-0,175)^2}} \tag{2.19}$$

Following this definition, SGP compliers distance from the decision boundary ranges from 0,00076 to 1,7305 whereas SGP non-compliers distance is between 0,0027 and 5,0717. We thus observe that some non-compliers are really far from the decision boundary as for example Ireland or Portugal in 2011, Greece from 2008 to 2013 or Slovenia in 2014. Such cases are really hard to help to run in the compliance subspace, Greece is the better example that was under strict European Commission monitoring for 10 years following the sovereign debt crisis.

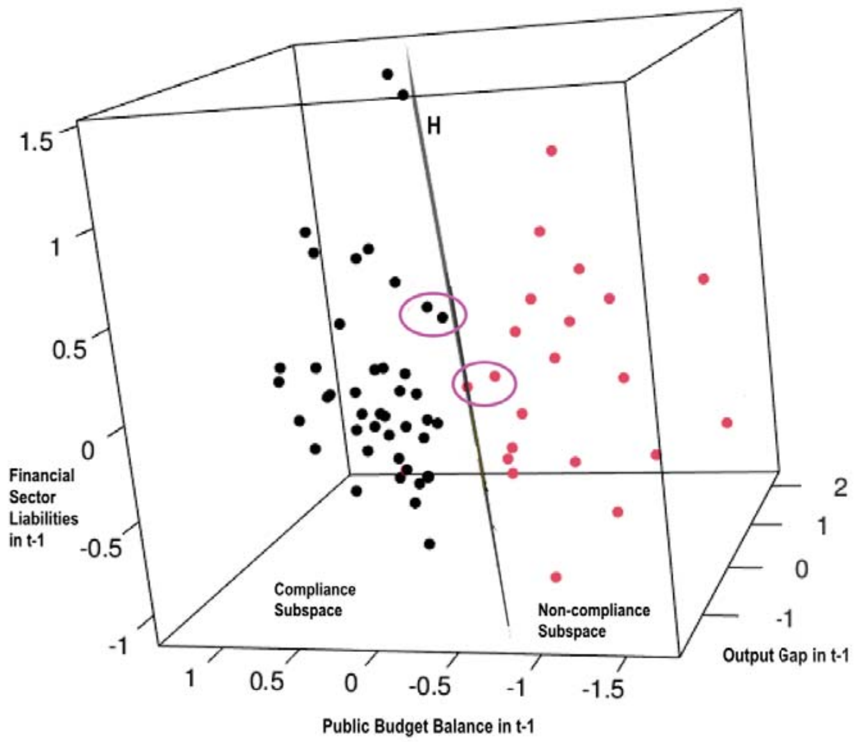



Figure 2.8: **Linear decision boundary in three Dimensions:**

Figure 8 shows $\frac{1}{5}$ of our dataset[29] and the linear separator hyperplane[30]. We can see that red dots circled in purple are the closest non-compliers from the Hyperplane. These observations could be easier influenced by policies to move in compliance area. These

[29] These observations are randomly selected. We do not present all the observations to make the figure clearly legible.

[30] In Figure 8 the linear separator only integrates three dimensions of our separator hyperplane H which is in 8 dimensions, and and it is therefore summed up to $\mathrm{H} : \alpha x + \beta y + \gamma z + b = 0$ .

points could correspond to Belgium in 2014 or Croatia in 2012 that present low distance from hyperplane. Black dots circled in purple also require attention since they are not so "far away from the cliff" and European Commission should monitor them. Such case could correspond to the Slovak Republic in 2007 for example.

## 2.6 Conclusion

While fiscal rules may achieve fiscal discipline and may be considered as performant according to chapter 1, this second chapter 2 extended the work on fiscal rules' performance. Indeed, fiscal discipline is a process that must be well-conduct from the introduction and use of tools such as fiscal rules and then it requieres a deep monitoring. Consequently, fiscal rules compliance is a major topic of fiscal discipline and needs several efforts to strenghten fiscal rules performance and let them properly operate.

This chapter 2 focused on the 28 EU member states over the period 2006-2018 and offered a new perspective in the debate which conflicts the dissuasive with the preventive arm: we propose a less complex but strong preventive arm, that will minimize the cases that ask for the dissuasive arm. Indeed, this methodology may serve as risk-management tool to design recommendations from the central authority to the "non-complier" countries. Instead of sanctions, the central authority may promptly (one year before) propose a set of well-targeted interventions that will change the fate of the "non-complier" (interventions that will change its position in the feature space: from the non-compliance subspace, to the compliance subspace). Such forecasting tool may be a solution face to the citation put at the begining of this chapter and that asks for political efforts "to be better aligned with rule compliance" (Caselli et al. [2018]).

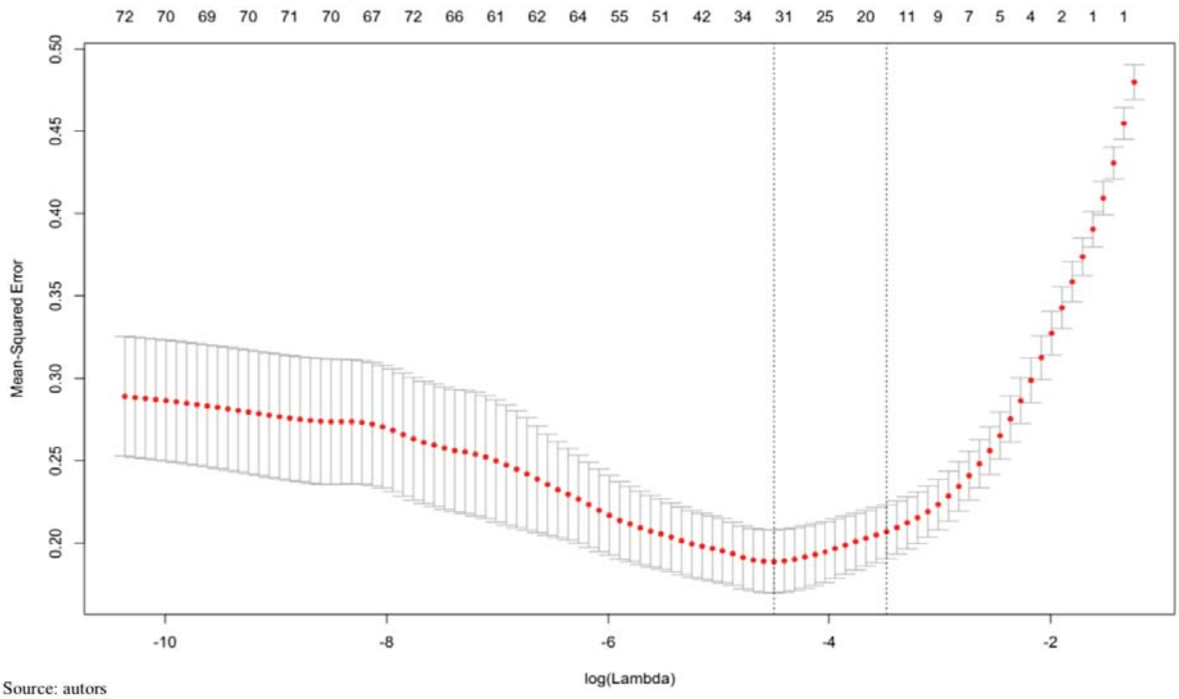
To build such tool, this chapter proposed a new Machine Learning based forecasting model on the compliance with the SGP for the EU member states. We focused our study on the public deficit rule, since a prompt forecasted of the 3% deficit limit can be fixed in a year. The same is not true for the public debt because when it is derailed, it needs multiannual recovery programs. A set of 8 features from a dataset of 141 variables is identified as key predictors thanks to the LASSO feature selection methodology. The chapter compared different Support Vector Machines models (three kernels: linear, quadratic and RBF), with a standard econometric approach as a benchmark, the Logit. The top performing model, trained in a K-fold cross validation set-up, used the linear kernel, and yielded

91.7% forecasting accuracy in the whole dataset (forecasting accurately 112 out of the 135 cases of non-compliance and 222 out of the 229 cases of compliance).

The findings may be examined under certain views. First, the chapter feeds the discussion about “The Impact of Machine Learning on Economics" (Athey [2018]). Indeed, the Machine Learning models provide high forecasting power, and they should be considered in fiscal policy outcome forecasting and risk events prevention. In our case the Machine Learning models outperformed in every case (except one) their Econometrics counterpart. Our study may open the way to the use of this type of models in other macroeconomic studies. Second, this chapter could be interpreted as a “risk-management approach” applied to fiscal surveillance and offers a solution to the need for fiscal framework simplification. Such simplification appears necessary for forecast endorsement by independent fiscal councils (Darvas et al. [2018], Debrun et al. [2019]). Our findings could lead to a first step in the European fiscal framework reform: i) MIP scoreboard indicators could be used in European Commission recommendations to help countries with their fiscal difficulties rather than for implementing excessive imbalance/deficit procedures; ii) simple advanced indicators could be implemented in a alert mechanism to prevent SGP deviations. iii) There are several possibilities for future research: i) the analysis could be conducted on the Compliance Tracker Database (Larch and Santacroce [2020]) that includes data compliance with the other fiscal rules included in the SGP, such as the structural balance rule and the expenditure rule; ii) the model could also be transposed to national fiscal policy outcomes forecasting using available dataset at national level; iii) these models could also be extended to other macroeconomic outcomes in forecasting the way to achieve monitoring objectives (as in monetary policy issues and macro prudential policies).

# Appendices

Appendix 1 : LASSO results




Source: autors

## Appendix 2. Descriptive Statistics

| Variables | Correspondance Variables | N | Mean | Min | Max | sd |
|---|---|---|---|---|---|---|
| Y | Dummy variable =1 if 3% limit was complied in t | 364 | 0.629 | 0.00 | 1.00 | 0.483 |
| X1 | Dummy variable reflecting if the country was an advanced country in t-p | 364 | 0.785 | 0.00 | 1.00 | 0.410 |
| X2 | Dummy variable reflecting if the country was an Emerging country in t-p | 364 | 0.214 | 0.00 | 1.00 | 0.410 |
| X3 | Dummy variable reflecting if the country was a Resource-rich country in t-p | 364 | 0.00 | 0.00 | 0.00 | 0.00 |
| X4 | Dummy variable reflecting if the country was an EU membership in t-p | 364 | 0.967 | 0.00 | 1.00 | 0.178 |
| X5 | Dummy variable reflecting if the country was a Federal Country in t-p | 364 | 0.107 | 0.00 | 1.00 | 0.309 |
| X6 | Dummy variable reflecting if the country was a eurozone member in t-p | 364 | 0.642 | 0.00 | 1.00 | 0.479 |
| X7 | Dummy variable reflecting if the country was submitted to an enforcement procedure related to the supranational fiscal rules in t-p | 364 | 0.967 | 0.00 | 1.00 | 0.178 |
| X8 | Gross domestic product, deflator, in t-p | 364 | 100.84 | 62.69 | 146.3 | 10.31 |
| X9 | Total investment in t-p | 364 | 22.47 | 9.819 | 41.53 | 4.756 |
| X10 | Gross national savingss t-p | 364 | 21.61 | 5.099 | 33.70 | 5.525 |
| X11 | Inflation, average consumer prices index, in t-p | 364 | 98.10 | 67.04 | 169.8 | 14.89 |
| X12 | Population in t-p | 364 | 17.92 | 0.403 | 82.66 | 22.63 |
| X13 | General government revenue in t-p | 364 | 41.96 | 25.94 | 56.36 | 6.546 |
| X14 | General government total expenditure in t-p | 364 | 42.63 | -7.824 | 65.047 | 11.19 |
| X15 | General government net lending/borrowing in t-p | 364 | -2.601 | -32.02 | 5.129 | 3.590 |
| X16 | General government gross debt in t-p | 364 | 60.41 | 3.664 | 183.4 | 34.96 |
| X17 | Net External Positions in t-p | 364 | -37.31 | -198.7 | 65.2 | 50.93 |
| X18 | Current account balance in t-p | 364 | -0.869 | -23.90 | 13.80 | 6.021 |
| X19 | Current account balance variations over 3 years in t-p | 364 | -1.214 | -21.00 | 11.00 | 5.722 |
| X20 | Real Effective Exchange Rate in t-p | 364 | 0.770 | -20.40 | 36.00 | 6.681 |
| X21 | Global export market share -% change over 5 years - in t-p | 364 | 4.547 | -31.68 | 95.57 | 23.88 |
| X22 | Nominal unit wage cost -% change over 3 years - in t-p | 364 | 6.948 | -21.00 | 78.30 | 10.35 |
| X23 | Debt of private sector in t-p, consolidated -% of GDP | 364 | 144.9 | 39.10 | 379.4 | 70.46 |
| X24 | Liabilities of the financial corporations sector, -% change over 1 year - in t-p | 364 | 8.145 | -17.60 | 115.6 | 12.85 |
| X25 | Unemployment rate - 3-year average - in t-p | 364 | 9.047 | 3.700 | 26.30 | 4.101 |
| X26 | Unemployment rate in t-p | 364 | 8.976 | 2.900 | 27.50 | 4.324 |
| X27 | Gross domestic product (real GDP) -% change over 1 year - in t-p | 364 | 1.966 | -14.80 | 25.10 | 3.837 |
| X28 | Gross fixed capital formation in t-p -% of GDP - | 364 | 21.93 | 11.50 | 37.40 | 4.196 |
| X29 | Gross domestic expenditure on R & D in t-p -% of GDP - | 352 | 1.504 | 0.370 | 3.750 | 0.877 |
| X30 | Direct investment in the reporting economy (flow) in t-p -% of GDP - | 364 | 25.01 | -264.1 | 1336.6 | 118.6 |
| X31 | Direct investment in the reporting economy (stocks) -% of GDP | 364 | 350.2 | 4.200 | 9479.1 | 1135.5 |
| X32 | Net trade balance of energy products in t-p -% of GDP - | 364 | -3.212 | -14.90 | 2.300 | 2.062 |
| X33 | Real effective exchange rate, Euro area trading partners -% change over 3 years | 364 | 1.487 | -21.70 | 38.90 | 6.335 |
| X34 | Terms of trade (goods and services) -% change over 5 years - in t-p | 364 | 1.102 | -10.20 | 28.30 | 4.870 |
| X35 | Market share of world exports, volumes -% change over 1 year - in t-p | 364 | 0.495 | -10.30 | 36.40 | 4.816 |
| X36 | Labor productivity -% change over 1 year - in t-p | 364 | 1.294 | -7.700 | 20.90 | 2.784 |
| X37 | Residential construction in t-p -% of GDP - | 350 | 4.347 | 0.600 | 13.50 | 2.121 |
| X38 | Employment -% change over 1 year - in t-p | 364 | 0.658 | -14.30 | 6.5 | 2.377 |
| X39 | Dummy variable reflecting if there is a Crisis in t-p | 364 | 0.307 | 0.000 | 1.000 | 0.462 |
| X40 | Output gap (production function approach) in t-p | 364 | 0.167 | -12.89 | 20.29 | 4.345 |
| X41 | Oil Prices in t-p | 364 | 73.76 | 43.29 | 99.67 | 19.84 |
| X42 | Bonds yield in t-p | 351 | 3.849 | 0.090 | 22.50 | 2.460 |
| X43 | Foreign currency long-term sovereign debt ratings, index from 1-21 , in t-p | 364 | 16.71 | 2.842 | 21.00 | 4.019 |
| X44 | Fiscal Space in t-p | 364 | 3.966 | -60.18 | 119.1 | 33.11 |
| X45 | Fiscal Rules Index (by European Commission) in t-p | 364 | 0.542 | -0.948 | 3.404 | 1.068 |
| X46 | Structural Breaks in t-p | 364 | 0.225 | 0.000 | 1.000 | 0.418 |
| X47 | Government fragmentation in t-p | 364 | 0.707 | 0.491 | 0.861 | 0.097 |

Note: Y is the Dependent variable. X are potential predictors tested in the feature selection step. All variables used as predictor are a p lagged of the variable. We report lag-1 in descriptive statistics to solve space and because lag-1 contains informations about lag-2 and lag-3 also tested in the chapter. Fiscal Space is measured as the difference between country public debt and EU median debt for each year.

# REFERENCES

## Bibliography

A. Annett. Social Fractionalization, Political Instability, and the Size of Government. *IMF Working Paper No. 00/82*, 2000. 96

S. Athey. The Impact of Machine Learning on Economics. *Chapter in NBER book The Economics of Artificial Intelligence: An Agenda*, pages 507–547, 2018. 116

A. Barbier-Gauchard, K. Baret, and A. Minea. National Fiscal Rules Adoption and Fiscal Discipline in the European Union . *Applied Economics*, 53(20):2337–2359, april 2021. 85, 88

M. Bergman, M.M. Hutchison, and S.E. Hougaard Jensen. Promoting sustainable public finances in the European Union: The role of fiscal rules and government efficiency. *European Journal of Political Economy*, 44:1–19, 2016. 85, 88

H. Bk and S. Maciejewski. The symmetry of demand and supply shocks in the European Monetary Union. *Bank i Kredyt*, 2017. 94

O. Blanchard, A. Leandro, and J. Zettelmeyer. Redesigning the EU Fiscal Rules: From Rules to Standards. *Mimeo, contribution presented to the 72nd Economic Policy Panel Meeting*, 2020. 88

European Fiscal Board. Assessment of the fiscal stance appropriate for the euro area in 2021. *Report*, 2020. 86

F. Caselli, L. Eyraud, A. Hodge, F. Diaz Kalan, Y. Kim, V. Lledó, S. Mbaye, A. Popescu, W. H. Reuter, J. Reynaud, E. Ture, and P. Wingender. Second-Generation Fiscal Rules : Balancing Simplicity, Flexibility, and Enforceability - Technical Background Papers . *Staff Discussion Notes No. 18/04*, 2018. 83, 85, 88, 115

H. W. J. Cheng and I. Pitterle. Towards a more comprehensive assessment of fiscal space. *DESA Working Paper No. 153, United Nations*, 2018. 96

European Commission. European Commission Fiscal Rules Database. 2017. 89

T. Cordes, T. Kinda, P. Muthoora, and A. Weber. Expenditure Rules: Effective Tools for Sound Fiscal Policy? *IMF Working Paper 15/29*, 2015. 89

C. Cortes and V. Vapnik. Support Vector Network. *Machine Learning*, 20:273–297, 1995. 102

J. Creel. Ranking Fiscal Policy Rules: the Golden Rule of Public Finance vs. the Stability and Growth Pact. *Documents de Travail de l'OFCE 2003-04*, 2003. 88

T. Dabán. A Second-Generation of Fiscal Rules for Latin America. *IMF Public Financial Management Blog*, 2011. 88

Z. Darvas, P. Martin, and X. Ragot. European Fiscal Rules Require a Major Overhaul. *Notes du conseil d'analyse économique*, 47:1–12, 2018. 85, 116

X. Debrun and L. Jonung. Under threat: Rules-based fiscal policy and how to preserve it. *European Journal of Political Economy*, 57:142–157, 2019. 84, 85, 87, 88, 94

X. Debrun, L. Moulin, A. Turrini, J. Ayuso i casals, and M. Kumar. Tied to the Mast? National Fiscal Rules in the European Union. *Economic Policy*, 23:297–362, 2008. 88, 89

X. Debrun, M. Gerard, and J. Harris. Fiscal Watchdogs and Sound Fiscal Policy: Is the Barking Loud Enough to Tame Politicians? *MPRA Paper No. 96683*, 2019. 90, 116

M. Delgado-Téllez, V. D. Lledó, and J. J. Pérez. On the Determinants of Fiscal Non-Compliance: An Empirical Analysis of Spain's Regions. *IMF Working Paper 17/5*, 2017. 85, 89

L. Eyraud, X. Debrun, A. Hodge, V. Lledó, and C. A Pattillo. Second-Generation Fiscal Rules : Balancing Simplicity, Flexibility, and Enforceability. *Staff Discussion Notes No. 18/04*, 2018. 85, 88, 94

D. Foremny. Sub-national deficits in European Countries: The impact of fiscal rules and tax autonomy. *European Journal of Political Economy*, 34:86–110, 2014. 85, 88

M. Frenkel and C. Nickel. How Symmetric Are the Shocks and the Shock Adjustment Dynamics Between the Euro Area and Central and Eastern European Countries? *Journal of Common Market Studies*, 2005. 94

J. Friedman, T. Hastie, and R. Tibshirani. Regularization for Generalized Linear Models via Coordinate Descent. *Department of Statistics Stanford University*, 2009. 104

P. Gogas and T. Papadimitriou. Machine Learning in Economics and Finance). *(ED) Computational Economics*, 2021. 104

P. Gogas, T. Papadimitriou, and E. Chrysanthidou. Yield Curve Point Triplets in Recession Forecasting. *International Finance*, pages 207–226, 2015. 90

P. Gogas, T. Papadimitriou, and A. Agrapetidou. Forecasting bank failures and stress testing: A machine learning approach. *International Journal of Forecasting*, 34:440–455, 2018. 90, 109

M. Guerguil, P. Mandon, and R. Tapsoba. Flexible fiscal rules and countercyclical fiscal policy. *Journal of Macroeconomics*, 52:189–220, 2017. 88

W. Härdle, Y-J. Lee, D. Schäfer, and Y-R. Yeh. Variable Selection and Oversampling in the use of Smooth Support Vector Machines for Predicting the Default Risk of Companies. *Journal of Forecasting*, 28:512–534, 2009. 90

S. Hauptmeier and C. Kamps. Debt rule design in theory and practice - the SGP's debt benchmark revisited. *Presentation at ECFIN workshop "Fiscal rules in Europe"*, 2020. 85

Z. Huang, H. Chen, C-J. Hsu, and W-H. Chen ans S. Wuc. Credit rating analysis with support vector machines and neural networks: a market comparative study. *Decision Support Systems*, 37:543–558, 2004. 91

IMF. Fiscal Rules—Anchoring Expectations for Sustainable Public Finances. *Fiscal Affairs Department, IMF Policy Paper*, 2009. 88

H. Ince and T. Trafalis. A Hybrid Model for Exchange Rate Prediction. *Decision Support Systems*, 42:1054–1052, 2006. 86

G. Kopits and S.A. Symansky. Fiscal Policy Rules . *Occasional Paper No. 162*, 1998. 84, 87, 88

M. Larch and S. Santacroce. Numerical compliance with EU fiscal rules: The compliance database of the Secretariat of the European Fiscal Board (Database). 2020. 85, 89, 90, 96, 116

M. Larch, E. Orseau, and W. van der Wielen. Do EU Fiscal Rules Support or Hinder Counter-Cyclical Fiscal Policy?). *Journal of International Money and Finance*, 2021. 88

M. Wafula Nandelenga and M. J. Ellyne. Fiscal Rules and the Compliance debate: Why do Countries adopt Rules and fail to comply. *ERSA Working Paper No 815*, 2020. 85, 86, 89

V. Plakandaras, T. Papadimitriou, and P. Gogas. Directional Forecasting in Financial Time Series using Support Vector Machines: The US/Euro Exchange Rate. *Journal of Computational Optimization in Economics and Finance*, 5:125–139, 2013. 86

W.H. Reuter. National numerical fiscal rules: not complied but still effective? *European Journal of Political Economy*, 39:67–81, 2015. 87, 88, 89, 90, 94

W.H. Reuter. When and why do countries break their national fiscal rules? *European Journal of Political Economy*, 57:125–141, 2019. 85, 86, 89, 96

C. D. Romer and D. H. Romer. Phillips Lecture – Why Some Times Are Different: Macroeconomic Policy and the Aftermath of Financial Crises. *Economica*, 85:1–40, 2018. 110

A. Sacchi and S. Salotti. The impact of national fiscal rules on the stabilisation function of fiscal policy. *European Journal of Political Economy*, 37:1–20, 2015. 85, 88

A. Schaechter, T. Kinda, and N. Budinaand A. Weber. Fiscal Rules in Response to the Crisis—Towards the 'Next-Generation' Rules. A New Dataset. *IMF Fiscal Affairs Department*, 2012. 89

A. Schick. Post-Crisis Fiscal Rules: Stabilising Public Finance while Responding to Economic Aftershocks. *OECD Journal on Budgeting*, 2010. 88

R. Tapsoba. Do National Numerical Fiscal Rules really shape fiscal behaviours in developing countries? A treatment effect evaluation. *Economic Modelling*, 29:1356–1369, 2012. 88

I. Velickovski and A. Stojkov. Is the European integration speeding up the economic convergence process of the Central and South-Eastern European countries? A shock perspective. *Empirica*, pages 287–321, 2014. 94

R. Wiese, R. Jong-A-Pin, and J. de Haan. Can successful fiscal adjustments only be achieved by spending cuts? *Economic Policy*, 54:145–166, 2018. 95

C. Wyplosz. Fiscal Rules: Theorical issues and historical experience. *NBER Working Paper No. 17884*, 2012. 89

# Chapter 3

# Side-effects of fiscal rules' performance on social welfare in OECD economies

"The poor track record of compliance has raised questions about the rules' capacity to ensure fiscal sustainability. Finally, efforts to achieve formal compliance with rules have also had undesirable side effects, encouraging creative accounting and the compression of public investment and social spending, particularly in emerging and developing economies." Eyraud et al. [2018]

## 3.1 Introduction

As discussed from the outset, the performance of fiscal rules relies on a number of factors, including fiscal rules compliance. However, there is no guarantee that conducting enforcement of fiscal rules will influence only fiscal discipline. It may also affect the rest of the economy, with the potential to cause a decline in social conditions for citizens and workers. While a government is under a budget constraint, it could restrain public expenditure and thus affect its public spending composition. For example, complying with a fiscal rules' target may lead a government to reduce social or health expenditure, which could, in turn, have negative consequences for inequalities and quality of life. The potential re-allocation of public expenditure to achieve fiscal rules compliance thus implies severe effects, justifying a thorough investigation. Consequently, fiscal rules' performance may come with side-effects and this chapter looks to address the following question: is fiscal rules' compliance detrimental to the economy and, in particular, for social welfare? To tackle this issue, this chapter considers the Budget Balance Rules' (BBR) compliance effects on macroeconomic indicators and social welfare proxy indicators in 16 countries between 2004 and 2015.

Social welfare is a broad concept with a seminal definition that covers basic human needs and originates from Maslow's pyramid (Maslow [1970]). Since the $18^{th}$ century, utilitarians such as Jeremy Bentham and John Stuart Mill have developed the argument that societies and governments should promote "The Greatest Good for the Greatest Number". This "Greatest Good" broadly refers to happiness and acceptable levels of health, income, and social conditions. The World Health Organization (WHO) extended the definition in the Ottawa Charter (1986) by considering social welfare as "a state of complete physical, mental and social well-being". Consequently, social welfare refers to different economic and social concerns that we will try to capture through different channels identified in the literature.

To study how fiscal rules' performace affects social welfare, instead of fiscal rules strength or fiscal rules presence effectiveness, we focus on fiscal rules' compliance effect on the social area. The marked effect of fiscal rules (usually proxied by the Fiscal Rules Strength Index (hereafter FRSI) or fiscal rules adoption) on public finance has been well documented[1], with some evidence that fiscal rules may affect public spending. In this chapter, we want to assess whether compliance also plays a role. However, the pandemic crisis (2020-2021) hit people, businesses and the public finance, leading to new social challenges. Indeed, Blundell et al. [2020] provided evidence on the pandemic period, highlighting impacts on employment and ability to work, investments and health. Consequently, we need to consider with caution the effect of fiscal rules performance on the social field, because a decision-maker who wants to restore sustainable public finance and adopts fiscal rules for their disciplining effect may neglect potential side-effects on economic growth and social welfare.

This focus on the effects of fiscal rules compliance necessitates a rigorous definition of compliance. The chapter considers two definitions of fiscal rules compliance. The simplest definition of compliance is a binary reflection of whether the fiscal rules did or did not meet the limit (as in Reuter [2019]), but compliance may also be considered in a more sophisticated form[2]. Fiscal rules often include escape clauses or exceptions[3], making the task of defining fiscal rules compliance more complex. In the presence of such escape clauses, it does not appear reasonable to consider that a country is a non-compliant if it exceeded the limit, but the escape clause was activated. In that sense, it is possible to define compliance as a situation where a country either presents a targeted indicator under (or equal to) the limit, or where the indicator is above the limit, but an escape clause is activated. In the latter case, the country is exceptionally authorized to deviate and should not be sanctioned. Not considering the presence of escape clauses could distort the results by introducing an error in the definition of the "public policy treatment" (the compliance), itself, and thus in its effect that we are trying to estimate.

We follow a multi-step approach to the empirical analysis, with the identification of fiscal rules' compliance determinants being the first step. The chapter provides an investigation of fiscal rules' compliance determinants considering existing studies that addressed this identification issue (see Reuter [2019], Delgado-Téllez et al. [2017] or Chapter

[1] See e.g. Debrun et al. [2008], Bergman et al. [2016], Tapsoba [2012], Combes et al. [2018] and literature developed in Chapter 1.

[2] See also for an alternative definition, Larch and Santacroce [2020] who explained the concepts to construct the European fiscal rules' compliance Tracker. In this database the European deficit rule is complied with if the public balance is superior to 3% or if the limit is exceeded, the deviation should be smaller than 0.5% of GDP and over only one year.

[3] See Chapter 2 for further details on escape clauses.

2 of this thesis for example). In our analysis we focus on national fiscal rules, and more specifically on Budget Balance Rules (BBR) compliance. We follow a similar approach to that adopted in Chapter 2 by identifying the main determinants of fiscal rules' compliance with Machine Learning methods that select the most prominent variables among many potential determinants. The second step is the Treatment Effect measurement. We expect that complied fiscal rules may have effects that non-complied fiscal rules could not have, in particular potential side-effects, on social welfare. This second step uses, as dependent variables, different channels through which fiscal rules compliance may affect social welfare between 2004 and 2015.

This Chapter 3 contributes to the literature in several ways.

Our approach first extends traditional assessment of fiscal rules performance by considering the fiscal rules compliance effect instead of fiscal rules effectiveness usually proxied by fiscal rules presence or strength. In that sense, we can measure the performance of fiscal rules with regards to the ultimate objective set out in the rules. Our study thus excludes problems associated with approaches using composite indices, such as FRSI, that are time in- variant.[4]. That being said, variables relating to fiscal rules characteristics (including FRSI), are considered in the present approach by evaluating if they are key predictors for Budget Balance Rules' (BBR) compliance in the first step of our methodology.

Second, our use of Double/Debiased Machine Learning (DML) treatment (Chernozhukov et al. [2017], Chernozhukov et al. [2018]) for fiscal discipline assessment is unprecedented and excludes biases that may arise in studies on fiscal rules performance, as discussed in Heinemann et al. [2018]. Indeed, Heinemann et al. [2018] noted that the majority of studies assessing the impact of fiscal rules on fiscal discipline is highly biased because endogeneity is not adequately controlled. The assessment of the fiscal rules performance effects employs numerous methodologies, including Instrumental Variable (IV), system-Generalized Method of Moments (sys-GMM) and propensity-score matching (PSM), as in Chapter 1. IV and sys-GMM performance is highly dependent on the choice of instruments (see Fajeau [2021] for discussion on instruments used in GMM models for economics studies; and Belloni et al. [2018] for a debiased GMM estimator that uses Machine Learning tools). On the other hand, propensity-scores are related to

[4]This implies that they do not consider the current numerical target and do not consider for macroeconomic country situation.

random assignment (meaning that conditional independence assumption must hold according to Rosenbaum and Rubin [1983]) which constitutes a strong constraint and assumption to ensure the robustness of the PSM approach. The algorithm we use is based on Norman orthogonality and is supported by strong asymptotic properties, thus generating a useful estimator for causal inference. DML estimation avoids reverse causality bias (which often occurs with standard econometrics) and reduces the potential omission bias since we can test a vast number of predictors.

Third, we include a proxy measure for “voter preferences” to increase the robustness of our analysis. This provides a significant value added among the existing literature on fiscal rules since previous studies based the robustness of their results on the assumption that voter preferences do not affect the results, and proposed many econometric robustness tests. Nevertheless, there is no certainty that these studies can control for omission bias and, in particular, the importance of voter preferences discussed by Wyplosz [2012].

Our main empirical findings concern both the identification of determinants of compliance as we first highlight that voter preferences are one of the key determinants for BBR's compliance; and the effects of compliance. The relevance we found of voter preferences for national BBR compliance suggests that studies dealing with fiscal rules performance issues should carefully account for Wyplosz [2012]'s bias. Then, we provided some evidence on BBR's compliance side-effects on social welfare. The negative consequences of strict compliance[5] operate through public spending composition, which mainly affect the redistribution function by reducing social expenditure. We also observe that BBR compliance increases inequalities. Governments seem to not operate a trade-off between economic objectives and BBR's compliance since we do not find a significant effect of strict compliance on GDP growth rate. However, a compliance definition which incorporates the presence of escape clauses may affect the results since we find a positive effect of compliance on economic growth after accounting for escape clauses. This implies that introducing flexibility in fiscal rules' compliance definition matters for economic health. Nevertheless, the negative impact on inequalities is not solved by relaxing the compliance definition and demands new reflections on fiscal rules design to carefully preserve public social spending.

[5]Strict compliance refers to the definition of compliance that only considers if a country met or not the limit of the BBR. It does not take account for flexibility by not considering the presence of escape clauses.

The rest of the chapter is structured as follows. Section 2 develops the literature review on fiscal rules compliance effects and social welfare channelse, Section 3 describes the data by insisting on national budget balance rules' compliance measurement and exposes the stylized facts. Section 4 presents the methodology, Section 5 reports the benchmark results and policy recommendations, Section 6 concludes the chapter.

## 3.2 Literature review on fiscal rules compliance effects and social welfare channels

### 3.2.1 The identification of social welfare channels

The goal of the chapter is to study the effect of fiscal rules' compliance on several channels that make the link with social welfare. After the seminal work of Arrow [1951], the concept of social welfare was formalized in economics and relies to political economy. That's being said, Hediger [2000] discussed government trade-offs among social, ecological, and economic objectives. By studying the link between fiscal rules' compliance and social welfare, we here implement a testing analysis of the potential government trade-off between fiscal performance (reflecting here by fiscal rules' compliance), social and economic objectives. Our main challenge is the identification of social welfare channels that may be concerned by the effects of fiscal rules compliance. Indeed, the list of social welfare determinants may refer to a lot of candidates such as the level of development, institutions (Acemoglu [2003]), fiscal policy (Gosh and Roy [2004]) and monetary policy (Lawler [2001]), the international trade (Samuelson [1938]), the financial development (Marini [2005]), geography (Smith [1974])... Consequently, social welfare may be linked with economic indicators as well as social indicators, and this chapter tries to identify the ones on which fiscal rules' compliance may have an impact.

First of all, social welfare may be linked with GDP growth as more wealth/ressources could increase well-being. But, it also depends on how these ressources are used and Midgley [1999] explained that social welfare may be driven by the distribution of resources generated by GDP growth. GDP growth may thus affect social welfare itself, but also through an undirect channel constituted by government performance. Indeed, government performance may increase during favorable economic periods which are supported by significant GDP growth rate, because governments may be less constrained. Nevertheless, the fiscal rules' compliance effect is ambiguous regarding both economic growth and government performance. It could lead to an increase in government effectiveness[6] as suggested by Larch et al. [2021] but may also imply a trade-off between fiscal rules' compliance and GDP growth objectives (Bohn and Inman [1996]). We will thus look at the effect of fiscal rules' compliance on GDP growth per capita and government performance alternatively measured by government effectiveness and government efficiency indices. Also, government performance concept is close to the nature of public spending that governments implement as Midgley [1999] explained that government may use positive return from GDP growth to implement social programs. This discussion relates to the Musgravian functions[7] that governments face. We thus should pay attention to the composition of public expenditure because they constitute a tool to conduct the redistribution function. If public sector conducts inefficient spending, public spending may be damageable for economies. On the other hand, the government size may support the economy and enhance social welfare. In that sense, if fiscal rules' compliance may affect public spending to ensure fiscal discipline, the indirect effect on economic growth as well as on social welfare is not clear. Blundell et al. [2011] investigated the link between fiscal rules and economic growth but there is no reference to the effect of the compliance. We precise our main interest in the compliance effect, not the presence or the rigor of fiscal rules, and we study a potential higher social cost due to compliance.

Also, social welfare is related to the level of public debt (see e.g Flodén [2001] or Aiyagari and McGrattan [1998])[8]. The level of public debt could also be linked to the redistributive government function and help people in smoothing their consumption (Burbidge [1983]). But growing public debt also leads to the common pool problem (Wyplosz [2012]) that may appear negative for future generations. We therefore are interested in the link between fiscal rules' compliance and public debt. But, fiscal rules are numerical constraint that must be complied in a year, it thus appears difficult to assess a long-run effect on the stock of public debt. Indeed, fiscal rules' compliance may easier affect public

[6] Larch et al. [2021] measured government effectiveness using the World Band index.

[7] Allocation; Stabilization (Stabilization power of fiscal rules was already studied by Sacchi and Salotti [2015] or Guerguil et al. [2017] who highlight that fiscal rules are able to stabilize GDP variations and public expenditures); Redistribution.

[8] Flodén [2001] showed that variations in public debt may enhance social welfare. Aiyagari and McGrattan [1998] studied the question of the optimal amount of public debt for social welfare in the US.

deficit which is a short-term flow variable than the total stock of debt accumulated over many years. Our first assessment of the relationship between public debt and fiscal rules' compliance will be studied through the effect of fiscal rules' compliance on public deficit that feeds public debt. On the other hand, if fiscal rules' compliance may be able to decrease public deficit, this may correspond to a positive effect on fiscal discipline. According to findings from Chapter 1, fiscal rules performance transit through financial market by sending a positive signal to financial markets leading to a decrease in the interest rate on public debt. It may consequently reduce the debt burden and gives governments more leeway. In that sense, we are interested on the effect of fiscal rules' compliance on both public balance and public debt interest rate.

Otherwise, we also should conduct further investigation on inequalities that may be a direct and undirect channel of social welfare. Inequalities may introduce a direct channel with social welfare because they refer to the quality-of-life conditions. On the other hand, inequalities as they may be viewed as an undirect channel. Kuznet [1955]'s curve described a non-linear relationship between the GDP growth and inequalities. In the first steps of development of the economies, GDP growth comes with an increase of inequalities. After achieving a sufficient level of economic development, the countries could then reduce social inequalities by redistributing the accumulated wealth. As developed in the previous paragraph, GDP growth is a channel of social welfare which appears also linked with inequalities. Consequently, inequalities may be first be affected by GDP growth and then, inequalities may affect social welfare. Inequalities thus represent a key but complex link with social welfare and the relationship between fiscal rules performance and inequalities is not obvious. Studying developing countries, Combes et al. [2019] found that Expenditure Rules increase inequalities while Budget Balance Rules and Debt Rules not; whereas Hartwig and Strum [2019] showed that fiscal rules increase inequality based on disposable income measures in the European Union. In line with these studies, we are interested in the side-effects of fiscal rules, but we focus on fiscal rules' compliance effects and we propose to assess the compliance impact on inequalities measured by proxy indicators including the Gini index computed by the World Bank.

### 3.2.2 Assessment of fiscal rules compliance effects

The world economic crises of the last decades challenged the fiscal rules compliance, but they also increased the debt unsustainability risk, raising the discussion on the relevance of fiscal rules for sustainability recovery. Consequently, the debate put the design of fiscal rules at the center. The definition of an ideal fiscal rule proposed by Kopits and Symansky [1998] introduced the concept of enforceability[9]. To make fiscal rules binding, sanctions can be included in the fiscal rules' design (as it is the case in the Stability and Growth Pact (SGP)[10]) and independent fiscal councils should be in charge of monitoring[11]. Compliance thus appears being a major concept when assessing fiscal rules performance.

In the existing literature on fiscal rules compliance, a large part is devoted on the compliance determinants. This literature was developed in the Chapter 2 that also discussed the works related to the forecast of fiscal rule compliance. Another part of this literature studies the government behavior face to the fiscal rules' compliance and its effect on the economic indicators. Reuter [2015], studying the dynamic of compliance showed that even if fiscal rules aren't comply, governments implement efforts to move close to the limit. This work was extended to emerging and developing countries; including both national and supranational rules in Caselli et al. [2018]. Similarly, Eyraud et al. [2018] highlighted the "magnet-effect" describing the trend of government to move close to the limit of fiscal rules. Such studies point out the benchmark status that the fiscal rules seem to have, suggesting that compliance seems to be a goal for governments. On the other hand, paying attention to this compliance which may sometimes be forced -in the sense that it goes against economic and fiscal impulse needs- also constitutes a topic for economic studies. We set our study in this strand of literature which focuses on fiscal rules compliance effects. The effects of fiscal rules performance on some of the channels of social welfare we discussed in 3.2.1, were addressed by the literature. Nevertheless, the studies do not necessarily consider compliance as the indicator for fiscal rules perfor-

[9] As defined by Kopits and Symansky [1998], the ideal fiscal rule should be simple regarding the target, clear, enforceable, consistent in the time, accompanied by an adequate fiscal framework.

[10] The beginnings of European fiscal rules enforceability come from the Maastricht Treaty (1992) with the excessive deficit procedure. The supranational rule in the EMU has been formalized in the SGP. Indeed, in the event of a recession of at least 2% of GDP, the European Commission then considers the economy in an exceptional situation, lifting the obligations to comply fiscal rules included in SGP.

[11] See Beetsma et al. [2018] for an assessment of fiscal councils' effect on governments commitment.

mance. Also, they do not cover different channels focusing in only one social welfare indicator and a particular attention is dedicated to inequalities measures. Larch et al. [2021] showed that EU supranational fiscal rules compliance reduce public debt and promote counter-cyclical fiscal policies. Since we are interested in the potential side-effect of compliance on social welfare, we extend this part of the literature by investigating the effect of national budget balance rules compliance on public finance indicators and public spending composition. Any change in the spending allocation and redistribution function of government implied by fiscal rules' compliance may lead to side-effect on social welfare. This also builds a bridge between fiscal rules' compliance effect and inequalities. The side-effects of fiscal rules simple presence on inequalities were already addressed by Combes et al. [2019] and Hartwig and Strum [2019]. Combes et al. [2019] found that BBR do not imply an increase in inequalities for developing countries while Hartwig and Strum [2019] found a positive effect of fiscal rules on inequalities in the EU. Despite the divergence between these results, they do not put a word on compliance effect. We thus extend these works by assessing if countries that comply with their national budget balance rules generate a side-effect on inequalities which are related to social welfare.

Our study thus extends existing literature by investigating the impact of fiscal rules compliance on different economic indicators to evaluate the presence of a potential government trade-off between economic objectives and fiscal rules compliance. Most importantly, we extend the assessment of fiscal rules performance effect to the effect of compliance on inequalities and other channels of social welfare that have not already been considered with their relationship to social welfare in the literature. All these channels are derived from the discussion proposed in in 3.2.1 and the data section 3.3.2 describes the measurement of these indicators. Moreover, our work comes with a causal Machine Learning estimator that discards reverse-causality such as overfitting bias, allowing for an interpretable treatment effect of fiscal rules compliance.

## 3.3 Data and stylized facts on national Budget Balance Rules' compliance and social welfare

This section presents the data, in particular the construction of the economies retained for the analysis, the compliance indicator as well as the list of potential determinants of national budget balance rules' compliance.

### 3.3.1 Measurement of fiscal rules' compliance

The construction of our dataset is driven by several constrains:

First, fiscal rules are defined as a numerical constrain set on public finance indicators (leading to Budget Balance Rules (BBR), Expenditure Rules (ER), Debt Rules (DR) and Revenue Rules (RR)). Different types of fiscal rules may mean *different effects* (See for heterogeneities of fiscal rules effect Debrun et al. [2008] or the Chapter 1 of this thesis). In that sense we must study the compliance by type of rule. The selected rules must be comparable to obtain a reasonable average treatment effect and thus must hold over the same period[12]. We finally identified sixteen countries who had a Budget Balance Rules over the same period, but we could not identify enough countries who applied the other types of rules on a same period. The study includes the following sixteen countries[13] which had a BBR between 2004 and 2015: Chile, Costa Rica, Denmark, Estonia, Finland, Germany, Hungary, Indonesia, Japan, Malaysia, New Zealand, Peru, Spain, Sweden, Switzerland, United Kingdom. Fourteen of these countries are OECD economies and the dataset was increased by two countries that were also under a BBR on the period 2004-2015. The two non-OECD countries are Indonesia and Peru which could not be neglected to avoid any selection bias due to a possible voluntary selection of only OECD members. All Budget Balance Rules and their target's definition come from IMF Fiscal Rules Database (Schaechter et al. [2012]) and targeted values' sources are developed in Appendix 1. Appendix 2 summarizes all BBR retained in this analysis and provides details on their definition.

Second, we had to precisely define each BBR regarding the possible presence of exclusion clauses. Because we first adopt a simple definition of compliance - i.e. a country complied with (resp. did not comply with) the BBR whether it presents an indicator above or equal to (resp. below) the target -, we must take into account the presence of escape clauses that allow countries to meet the limit during "exceptional" economic circuntances

[12]We could skew the distribution of the sample by taking countries that have had a fiscal rule for 5-year and compare them to countries that had a fiscal rule throughout our study period.

[13]Despite Israel also had a BBR all over this period, it is discarded due to the annual change in the targeted value of BBR which does not match with the definition of an annual numerical target.

[14]. The presence of escape clauses can disrupt the distribution of compliance as they are a part of the fiscal rules' design. The escape clauses also set a huge debate on the compliance definition that we try to consider by testing the influence of such escape clauses on our results. Our robustness tests regarding escapes clauses are two-fold: i) we test whether the presence of an escape clause is a key determinant for national BBR's compliance; ii) we conduct a robustness test of the treatment effect by removing all observations that did not complied with BBR that are designed with escape clauses[15].

Third, some countries of our dataset need special attention. (1) United Kingdom abandoned its Golden rule in 2009 due to the Global Financial Crisis (GFC) that led to an excessive deficit rendering the compliance with the Budget Balance Rule impossible. We assume that United Kingdom (UK) voluntarily did not comply the Golden rule in 2009 and treat the UK as a non-complier in 2009. The UK introduced a new Budget Balance Rule in 2010 which targets a balanced structural budget at the end of 5 years (2014). This new BBR is interpretated as an annual change targeted variables (Caselli et al. [2018], Reuter [2019]). We then verify if this assumption does not affect our results by then conducting a robustness test which consists in removing this year-corresponding-observation from our sample. (2) Hungary had two fiscal rules between 2009 and 2011. Only the BBR that concerned General government is considered since all other countries are treated with only one BBR. Also, Hungary seems to stop having fiscal rules after 2011 in the IMF Database (Schaechter et al. [2012]). But the Fiscal Compact (also known as "The Treaty on Stability, Coordination and Governance (TSCG)") was transposed in EU members' national law. In that sense we could not consider that there is no BBR applied at national level in Hungary. We thus assume that structural deficit should be above 0.5% (because debt is higher than 60%; as described in TSCG). We also conduct a robustness test that consists in removing Hungary observations after 2011 to give a proof that our results are not sensitive to this interpretation. (3) In Caselli et al. [2018] the Golden rule of Japan is considered only between 1990 and 1993 because waiver looks as requested since 1975. However, this rule is well considered in the IMF database and we assume that this is a voluntary attitude of Japan regarding its rule. Japan has never complied with its rule over the study period, but

[14] For example, the European Commission defines exceptional circumstances in the SGP escape clauses as a recession of 2% of GDP.

[15] Such observations may be interpretated as compliers if they are allowed to exceptionally deviate from their national rule. In that sense, we must control if including them as non-compliers following a simple definition of compliance, doesn't affect the results.

it is an assumed deviant behavior from Japan and we cannot neglect this aspect.

### 3.3.2 The potential determinants of national Budget Balance Rules' compliance and proxy variables for social welfare channels

First, Table 1 reports the dependent variables of our interest. According to the literature review (see section 3.2.1), we identified several channels related to social welfare which are named "social welfare related indicators" in table 1. We consider them as reasonable proxies for social welfare. Some of these channels are represented by macroeconomic variables as public balance, interest payments on public debt, general government gross fixed capital formation, general government final consumption[16], GDP per capita annual growth.

We study the effect of the national BBR compliance on GDP per capita growth to evaluate if governments can prioritise compliance over economic objectives, which could be detrimental to the economy as well as to social welfare. Furthermore, we expect that, when a government faces its BBR constraint, it will operate a change in its spending composition. This change may affect their expected return-effect on economic growth. For example, according to the economic theory, public infrastructure expenditure may be positive for economic growth whereas consumption expenditure not (Everaert et al. [2015]). In that sense, a government that complied its BBR may choose to favor public GFCF while decreasing social expenditure and thus, expects GDP growth in return. We thus need to evaluate the impact of BBR compliance on government expectations. To do so, we produced a measure for the GDP growth expectation based on a 5 years moving-average of the GDP growth.

On the other hand, the indirect effects of GDP on social welfare are mediated, as explained in the literature review, by the way resources are used. Thus we are also interested in the effects of compliance on the government performance and the composition of public spending. To assess the compliance effect on public spending composition we focus our attention on general government gross fixed capital formation as it may be considered

[16] General Government final consumption is divided in Government individual consumption (P.31 in Eurostat classification) which includes social transfers and government non-market production of individual goods and services (D.631 and D.632), and Government collective final consumption (P.32 in Eurostat classification) which concludes Government collective non-market output, other related to collective goods and services (P.132-5.631).

as productive spending (and thus potentially be beneficial for social welfare), and general government final consumption as it includes social spending (such as social transfers) which should promote social welfare. Government performance and its redistribution function may also be affected by fiscal rules performance because they may affect public finance indicators as well as other macroeconomic indicators (see the general introduction and chapter 1 for further developments) as governments that spend better and/or more efficiently are likely to improve the welfare of their citizens. Thus, dependent variables also concern government performance by including the Government Effectiveness Index from the World Bank, and our own constructed index of Government Efficiency which summarized the government Musgravian functions. We aim at comparing the effect of BBR's compliance on government Effectiveness and government Efficiency that are two different concepts[17]. Following Afonso et al. [2006] and Afonso et al. [2019], we construct a measure for Government Efficiency index computed over-year. We choose 3-over-years computation (instead of 5 years as often found in the literature) to reduce the time-invariance of the indicator. We use mean-min function to aggregate 3 sub-indicators which correspond to the Musgravian functions (see Afonso et al. [2006] or Afonso et al. [2019] for similar proxies): - the proxy for the distribution function is the Gini index; - the proxy for the stabilization function is constructed by the sub-aggregation of the GDP per capita growth rate and inflation (3 years average); - the proxy for the economic performance function is the unemployment.

Finally, due to the potential link between government performance and inequalities previously discussed in the literature review, we also introduced inequalities related measures among the dependent variables. These indicators are summarized by the Gini index from World Bank and the Poverty headcount ratio at 1.90$ a day which is defined as the percentage of the population living with less than 1.90$ per day.

Second, Table 1 summarized the list of potential predictors that may affect both the BBR's compliance and the dependent variables. In line with many results from studies analyzing the determinants of fiscal rules' compliance[18], we expect that the compliance will be affected by many macroeconomic environment variables named “Macroeconomic Environment Variables” in Table 1, but also by political variables (as the presence of election) named “Countries characteristic Variables” or variables related to fiscal rules' design

[17] See e.g. general introduction for discussion.

[18] Reuter [2019], Delgado-Téllez et al. [2017], Larch et al. [2021] for example

(as the strength of fiscal rules) named "Fiscal Rule Related characteristics". The justification for evaluating the importance of these potential predictors is similar to the motivations of Chapter 1 that implemented robustness tests on the effect of political factors as well as factors linked with fiscal rules. Chapter 2 also introduced political variables as the government fragmentation (in opposite to government stability) as well as rules characteristics (such as the Dummy variable reflecting if there is an enforcement procedure). Political variables can interfere with the governments behavior, especially with regard to their compliance with fiscal rules. As illustration, the government stability or the rule of law index (that assesses the extent which economies adhere to the rule of law in practice) reflect political credibility which may play a role in governments commitment and their ability to fulfill their objectives. On the other hand, characteristics of fiscal rules (which refer to the ones we described in the introduction such as the coverage level) could also affect the credibility of the rule itself and thus should be considered when assessing fiscal rules' performance.

Among these list of potential determinants, we are interested in finding those which are recurrent from one country to another and contain useful information to explain the compliance with the budget balance rules.

To extend the list of potential determinants and improve the empirical literature on fiscal rules' compliance's determinants, we follow Debrun and Kumar [2007] and Wyplosz [2012] who suggested that fiscal rules effect could suffer from reverse causality bias. This argument is also supported by recent findings in Heinemann et al. [2018]. Such bias may still hold when assessing fiscal rules' compliance effect. Indeed, if compliance could imply differences in macroeconomic indicators, these ones could also influence the governments in their commitment (degraded public finance can strengthen the governments' willingness to comply with fiscal rules in order to restore sound public finance). We will thus be really careful in the use of lagged macro variables in the tested dataset for potential predictors. Moreover, Wyplosz [2012] argued that Voters' Preferences may affect government behaviors, especially regarding the fiscal rules' compliance. Indeed, decision-makers may be tempted to break fiscal rules aiming at increase social spending to be re-elected. Conversely, if voters prefer disciplined governments, public authorities could force compliance with the rules. We thus follow Funk and Gathmann [2013] that used Latent Factor analysis to compute a measure of voter preferences for Swiss Canton. To do so, use five variables that reflect the voter behavior namely Unemployment, Age dependency ratio (old in % of working-age population), the share of votes obtained by the largest government party, the vote share obtained by the first opposition party, the vote share obtained by independent parties. The Chi-test revealed (for varimax and promax rotation) that 2 factors are sufficient. We will thus use these two factors as control variables since they constitute good proxies for voter preferences[19].

[19]If the feature selection step reveals that one or both factors are a key determinant for fiscal rules' compliance, it will give an empirical recommendation for studies on fiscal compliance to control for voter preferences.

Table 3.1: **Variables Overview**

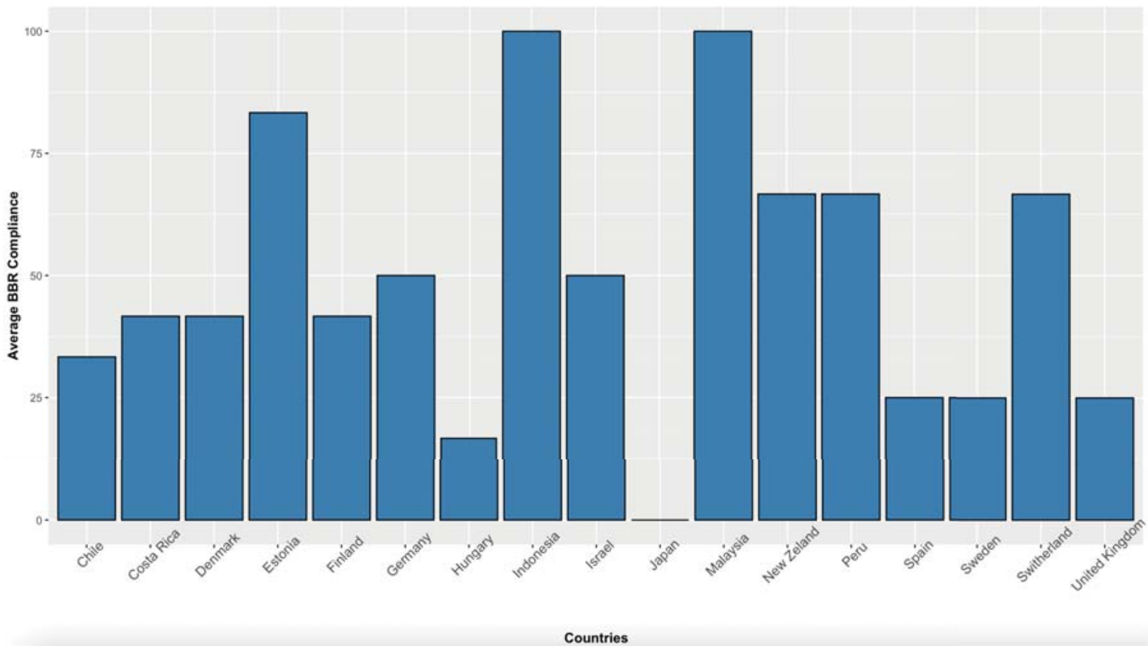
| Variables | Correspondence Variables | Source/Database | |
|---|---|---|---|
| dependent | Public Balance (in % of GDP) | World Bank | Social Welfare Related Indicators |
| Dependent | Interest payments (in % of expense) | World Bank | |
| Dependent | GG Gross Fixed Capital Formation (in % of GDP) | World Bank | |
| Dependent | GG Total Spending (in % of GDP) | World Bank | |
| Dependent | General Government Final Consumption (in % of GDP) | World Bank | |
| Dependent | GDP per capita expectations | | |
| Dependent | GDP per capita (annual growth) in $t+1$ | | |
| dependent | Government Effectiveness Index | World Bank | |
| dependent | Government Efficiency Index | Author's calculation | |
| Dependent | Gini index | World Bank | |
| Dependent | Poverty headcount ratio at 1,90$ a day (2011 PPP) (% of population) | World Bank | |
| Predictor | Control of corruption | WWGI | Countries Characteristic indicators |
| Predictor | Political Stability | WWGI | |
| Predictor | Regulatory Quality | WWGI | |
| Predictor | Rule of law | WWGI | |
| Predictor | Voice and Accountability | WWGI | |
| Predictor | Dummy variable reflecting if the country is an Advanced country | IMF Fiscal rules' Database | |
| Predictor | Dummy variable reflecting if the country is a Resource Rich country | IMF Fiscal rules' Database | |
| Predictor | Dummy variable reflecting if the country is an Emerging country | IMF Fiscal rules' Database | |
| Predictor | Dummy variable reflecting if the country is an Advanced country | IMF Fiscal rules' Database | |
| Predictor | Dummy variable reflecting if the country is a EU member | IMF Fiscal rules' Database | |
| Predictor | Dummy variable reflecting if the country is member of a currency union | IMF Fiscal rules' Database | |
| Predictor | Political system | WWGI | |
| Predictor | Dummy variable reflecting if there was an legislative election in this year | WWGI | |
| Predictor | Dummy reflecting if there was an executive election in this year | WWGI | |
| Predictor | Executive Index of Electoral Competition | WWGI | |
| Predictor | The number of years the chief execute has been in place | WWGI | |
| Predictor | Time since formation of the largest government party | WWGI | |
| Predictor | Proxy 1 for Voter's preferences | Author's calculations with LFA | |
| Predictor | Proxy 2 for Voter's preferences | Author's calculations with LFA | |
| Predictor | Well specified escape clauses | IMF fiscal rules' Database | Fiscal rule Related characteristics |
| Predictor | Monitoring of compliance outside government | IMF fiscal rules' Database | |
| Predictor | Formal enforcement procedure | IMF fiscal rules' Database | |
| Predictor | Coverage level | IMF fiscal rules' Database | |
| Predictor | Dummy variable reflecting if an independent body sets budget assumptions | IMF fiscal rules' Database | |
| Predictor | Dummy variable reflecting of an independent body monitors implementation | IMF fiscal rules' Database | |
| Predictor | Dummy variable reflecting if the BBR is a Golden rule | Author's narrative approach and IMF fiscal rules Database | |
| Predictor | Dummy variable for economy conjuncture | | Macroeconomic Environment Variables |
| Predictor | Oils rents | | |
| Predictor | Interest payments on debt in $t-1$ | | |
| Predictor | Gross Fixed Capital Formation (annual growth) in $t-1$ | | |
| Predictor | Gross Fixed Capital Formation (in % of GDP) in $t-1$ | | |
| Predictor | The Current account balance in $t-1$ | | |
| Predictor | The Unemployment rate in $t-1$ | | |
| Predictor | Trade (in % of GDP) in $t-1$ | | |
| Predictor | Inflation, consumer prices (annual %) in $t-1$ | | |
| Predictor | Inflation, GDP deflator (annual %) in $t-1$ | | |
| Predictor | Wage in $t-1$ | | |
| Predictor | GDP per capita growth (annual %) in $t-1$ | | |
| Predictor | Labor Force in $t-1$ | | |
| Predictor | External Balance in $t-1$ | | |
| Predictor | General Government budget balance in $t-1$ | | |
| Predictor | General Government final consumption in $t-1$ | | |
| Predictor | Central government debt (in % of GDP) in $t-1$ | | |
| Predictor | Gross savings in $t-1$ | | |
| Predictor | Total expenses in $t-1$ | | |

Note: GG = General Government; LFA = Latent Factor Analysis; GDP per capita expectation is computed using a 5 years moving-average approach based on GDP per capita data coming from the World Bank.

### 3.3.3 National BBR compliance and social welfare stylized facts

This part aims at illustrating the intuitions regarding the potential links between fiscal rules' compliance and social welfare channels.

Figure 3.1 first shows a high heterogeneity in government behaviors regarding national BBR's compliance. While some countries as Estonia, Indonesia, Malaysia, or Switzerland take care of the compliance, other as Japan, Hungary or Spain exhibit a poor compliance record. These countries are historically, socially, and structurally different. In that sense, we expect that the identification of the key common determinants for the BRR's compliance to help us to provide explanations about such differences across countries' compliance record.




Note: "0" means that the country never complied with its national BBR. "100" means that the country complied every year across 2004-2015 period.
Source: Author.

Figure 3.1: **Average Budget Balance Rules' (BBR) compliance between 2004 and 2015, in %**

Face to this heterogeneity between countries regarding the BBR's compliance, we are interested on the potential effects of these differences on the economy and social welfare. We thus propose a graphic comparison of the compliers group (countries that complied with their BBR) over the non-compliers group (countries that did not comply with their BBR). We are interested in the analysis of the social welfare related indicators of these two groups by comparing the median of each group. Appendix 3 provides a comparison of public spending and Gini index between each group by quantiles.

In Figure 3.2, the median of total public expenditure (in % of GDP) looks higher in countries that did not comply with their BBR. It suggests that countries from the compliers group operate a cut in some part of their public spending to comply with their BBR. This fact seems to reflect the disciplining effect of compliance since the compliers implement more efforts by reducing total public spending to comply with their national BBR. Nevertheless, this simple overview does not provide information on which type of public spending are affected by the cut from the compliers. Among public spending we may find unproductive spending such as interest payment on public debt or productive investment such as public GFCF. Otherwise, social spending, such as transfers, are included in the government Final Consumption expenditure which are a part of total public expenditure. We thus need a deep empirical analysis of the effect of BBR's compliance on public spending composition.

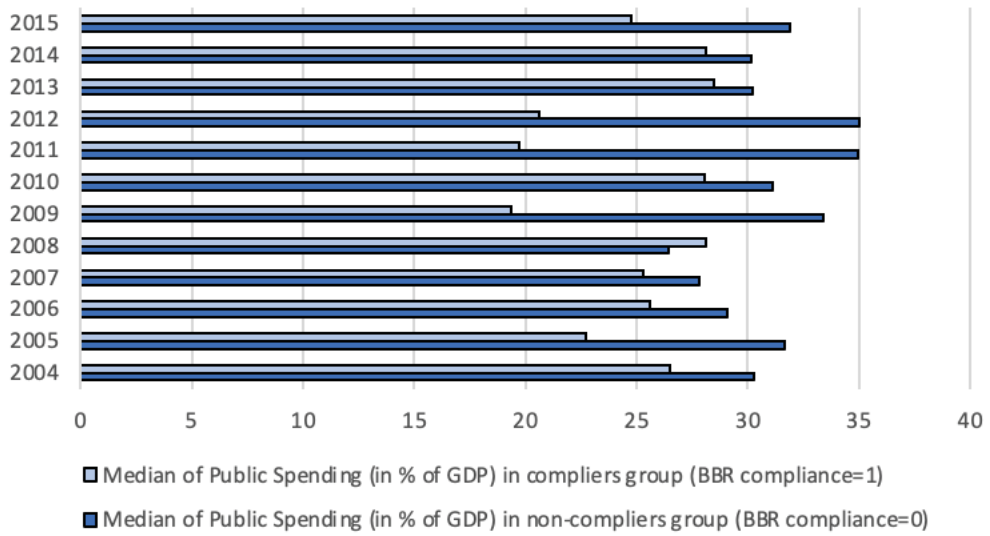
In parallel, figure 3.3 shows that the median of the Gini index seems to be higher for the BBR-compliers which suggests that inequalities are higher for them. A possible way to link these graphical findings is that the cut in public spending seems to be done through public social spending and thus need a careful attention. We also see that the differences in Gini index highly increased after the Global Financial Crisis (GFC) as the Gini index median became even more higher for compliers, suggesting that the GFC increased the social costs for compliance.

We must note that in the year of the GFC shock (2008) we observe the opposite to what we described above for the rest of the study period. Indeed, the Gini index is lower for compliers, while the total public expenditure is higher than non-compliers. We see two possible explanations:

i) the exceptional circumstances generate exceptional facts. It may be due to the escape clauses application in this year which means that there was no BBR enforcement letting countries to implement their fiscal impulsion to help in economic recovery. In such conditions, the distinction between "compliers" and "non-complier" is no longer so clear. Finally, in times of crisis, few countries comply with their rule (in the strict sense/without taking into account the escape clauses) and, in general, public spending increases to support activity. On the other hand, the deterioration of economic conditions, in particular employment, also increases social inequalities. When we move away from the crisis shock, we observe that countries which tend to comply with their BBR seem to spend less and exhibit more inequalities;

ii) it is also possible that countries that are used respecting their rules will be in better shape when the crisis arrives. This would give them more scope to limit the crisis (less inequality and more public spending). But when conditions return to normal, more than half of the countries that respect their rules have higher inequality and lower spending again.

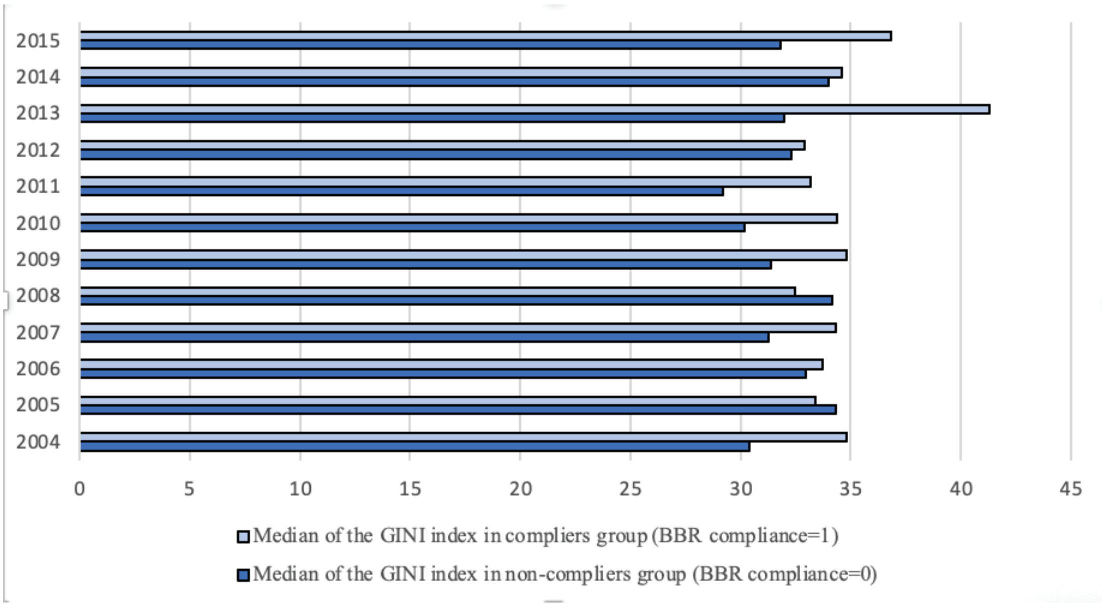
This point launches the importance of analyzing the definition of compliance, in particular a flexible definition that incorporates escape clauses.




Source: Author.

Note: *BBR compliance* is a Dummy variable taking the value 1 if the country complied with it BBR and value 0 if it did not comply with. The compliance definition considered here is the simplest one that does not consider flexibility and escape clauses in fiscal rules' design. The sample covers our sixteen studied countries.

Figure 3.2: **Comparison of the median of the Public Spending between BBR compliers and BBR non compliers between 2004 and 2015**




Source: Author.

Note: *BBR compliance* is a Dummy variable taking the value 1 if the country complied with it BBR and value 0 if it did not comply with. The compliance definition considered here is the simplest one that does not consider flexibility and escape clauses in fiscal rules' design. The sample covers our sixteen studied countries.

Figure 3.3: **Comparison of the median of the Gini index between BBR compliers and BBR non compliers between 2004 and 2015**

## 3.4 Methodology: Feature selection and Double/Debiased Machine Learning estimator

### 3.4.1 Treatment Effect Estimation

Recently, some studies focused on the usefulness of Machine Learning (ML) on the causal inference that belongs to the applied econometric field (Varian [2014], Mullainathan and Spiess [2017] or Athey and Imbens [2017]). Several techniques were developed to improve ML performance in the work of the treatment effect methodology. Among these techniques we can find: i) sample splitting which uses different data partition to select the best models and parameters (see Athey et al. [2016] or Wager and Athey and Imbens [2017]) and ii) orthogonalization (e.g. Chernozhukov et al. [2017]). Such approaches imply properties as asymptotic normality in these ML estimators (see Athey et al. [2017] for the general semiparametric case or Chernozhukov et al. [2018] for the average treatment effect case).

The main goal of our work is to estimate confidence intervals for a low-dimensional parameter $\beta_0$ with high-dimensional nuisance parameter $\eta_0$. The $\eta_0$ parameter should be estimated with the recent nonparametric statistical methods belonging to the Machine Learning (ML) field. ML methods highlight high level forecasting power (see Chapter 2 of thesis or Härdle et al. [2009] and Gogas et al. [2018]). However, this performance in forecasting does not imply inference performance for "causal" parameters. To solve such problem, Chernozhukov et al. [2017] developed Double/Debiased Machine Learning methodology (also called orthogonalized ML), introducing an approach inspired from Frisch-Waugh-Lovell (Frisch and Waugh [1933], Lovell [1963]) with a combination of feature selection and sample splitting aiming at proposing a strong estimator for causal parameters.

Our model is a partially linear model that could be written as:

$$Y = \beta_0 * D + \gamma_0(Z) + U, \qquad \mathbb{E}[U|Z, D] = 0, \tag{3.1}$$

with Y the outcome variable, D the treatment/policy variable, Z is a high-dimensional vector of controls/confounders, $\beta_0$ is our parameter of interest.
Z corresponds to control variables on the sense that the treatment D is defined as

$D = b_0 + \theta_0(Z) + V$ with $\theta_0 \neq 0$

If conditional exogeneity (view Rosenbaum and Rubin [1983]) is respected, $\beta_0$ corresponds to the average treatment effect of the treatment. The Double/Debiased Machine Learning (DML) works in several steps:

1) In a first step we will use two Machine Learning approaches[20] to predict Y and D on Z to obtain $\widehat{E[Y|Z]}$ and $\widehat{E[D|Z]}$. This step corresponds to the feature selection.

2) We then extract residuals $\widehat{W} = Y - \widehat{E[Y|Z]}$ and $\widehat{V} = D - \widehat{E[D|Z]}$.

3) Following Frisch-Waugh-Lovell procedure (Frisch and Waugh [1933], Lovell [1963]) we regress $\widehat{W}$ on $\widehat{V}$ that allows us to obtain $\widehat{\beta_0}$. This step is the orthogonalization procedure.

[20] Least Absolute Shrinkage and Selection Operator (LASSO) and the $l2$-boosting.

### 3.4.2 Feature Selection Estimators

Following Chernozhukov et al. [2017] and Chernozhukov et al. [2018], we will use different feature selection procedures as robustness tests that allow us to make our results generalizable. As techniques, we propose the Least Absolute Shrinkage and Selection Operator (LASSO) which was already presented in chapter 2 and the $l2-boosting$.

In the context of our analysis, we should keep in mind that the dependent variables of interest are continuous while the treatment effect (BBR's compliance) is a binary variable. In that sense, the following algorithms will be adapted of each case (continuous or binary). Since our main dependent variables (the overall public balance, the interest payments, the total public spending, the government final consumption, the GDP per capita expectation, the GDP per capita in $t+1$, the government Effectiveness, the Musgravian Index, the Gini index and the poverty headcount ratio) are continuous, we are able to report the efficiency using the Root-Mean-Squared-Errors of each feature selection model in the tables of results. Appendices 4 and 5 provide an illustration of fitted values distribution (for one of our variables of interest[21]) resulting from both feature selection algorithm and highlight the normal properties that allow such procedures.

**Presentation of the $l2$-BOOSTING**

The so-called Gradient Boosting is a Machine Learning application of Boosting which is based on sequential Ensemble. Ensemble learning method uses several learners to provide a final stronger learner. In that sense Boosting is an Ensemble technique that will produce several weak leaners used to construct a strong next learner that minimizes the total model prediction error. The weak learners (also named weak rules) are obtained by using ML algorithms on different distributions of our dataset.

[21] All fitted values distribution for all our variables of interest are available upon request to the author.

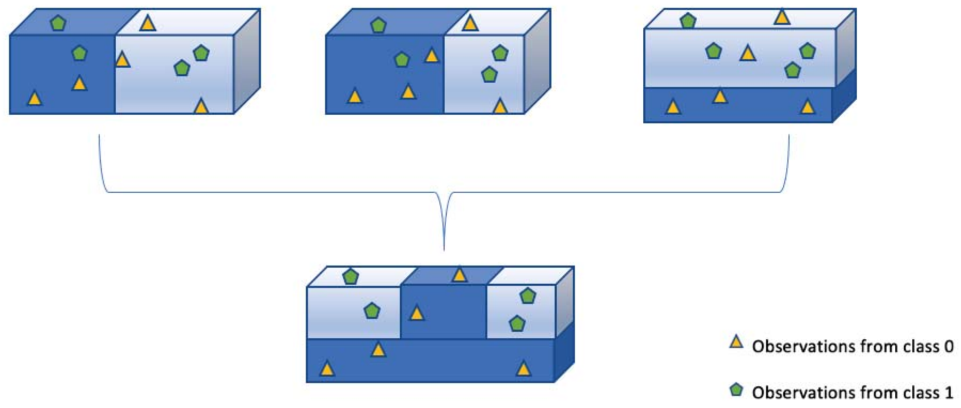



Figure 3.4: **Illustration of Boosting Algorithm**

The Figure 3.4 provides a simple illustration of how the Boosting algorithm works. In the first step, the algorithm analyzes the dataset and assigns equal weights to each sample. The false predicted observations provided by the "base" learner are identified in the second step. In the next iteration, the false predicted observations will be assigned to the next base learner with a higher weight. The algorithm continues with the repeats the weights update and forecasting until the ending criteria is met.

By definition, Gradient Boosting sequentially generates base learners that are more effective than the previous one. Gradient Boosting makes the overall model improving sequentially with each iteration.

Gradient Boosting optimizes the loss function of the previous learner. To do so, Gradient boosting adds a new model that adds weak learners aiming at reducing the loss function in order to overcome the errors in the previous learner's predictions.

The Boosting with $l2$-loss function follows the functional gradient descent procedure, including a $l2$-penalty term. Such procedures need an initialization step, by setting target outcomes for the first model. This algorithm is equivalent to the functional gradient descent technique. The main goal is to estimate the function:

$F : \mathbb{R}^d \longmapsto \mathbb{R}$, minimizing an expected cost

$$\mathbb{E}[C(Y, F(X))], C(.,.) : \mathbb{R} \times \mathbb{R} \longmapsto \mathbb{R}^+ \qquad (1)$$

where $Y_i$ is our dependent variable and $X_i$ the potential predictors for observations $i = 1, ....n$. When Y is continuous, the problem is solved through regression; when Y is discrete, we are in a classification issue. Cost function C(.,.) verifies important properties to make sure that gradient approach works well: it is smooth and convex in the second argument.

L2-Boost cost function is: $C(y, f) = \frac{|y-f|^2}{2}$ with $y \in \mathbb{R}$ or $y \in \{0,1\}$, $f \in \mathbb{R}$

Following Friedman et al. [2000], the population minimizers to estimate (1) is:

$$F(x) = \mathbb{E}[Y|X = x]$$

The application of functional gradient descent to the dataset lead to the minimization of the empirical risk and the estimation of F(.) given by:

$$n^{-1} \sum_{i=1}^{n} C(Y_i, F(X_i))$$

We thus apply this algorithm in a binary/classification issue when the dependent variable is the treatment (BBR (non-)compliance) which corresponds to the compliance determinants identification step. Then, we apply this algorithm in a linear approach for our main variables of interests (GDP growth, Government Spending and social indicators) that are continuous. For further details on Generic functional gradient descent and L2-boosting with linear/classification learners, see Bühlmann and Yu [2003].

## 3.5 Results and policy recommendations

This section develops the findings provided by our Double/Debiased Machine Learning (DML) estimator. Results first report the findings from the feature selection step. This step is crucial because it extracts information from both dependent variables and treatment (namely BBR compliance) before assessing the treatment effect. We focus our attention to the identification of the determinants of national BBR's compliance because it is the retained indicator of fiscal performance of interest. We do not report the variables selected as determinants for the dependent variables. If any determinant of the BBR affects one or several of our dependent variables, this information is considered by our methodology developed in Section 3.4.1. The second part of the result presents the Average Treatment Effect (ATE) of the BBR's compliance on the dependent variables defined in section 3.3.2 which correspond to the social welfare channels.

### 3.5.1 Results from Feature Selection procedures

Table 3.2 reports the 10 key common determinants of the BBR's compliance retained by our two feature selection algorithms: the Dummy variable for economic crisis, the

Dummy variable reflecting the presence (or not) of escape clauses, the Dummy variable reflecting the presence (or not) of a formal enforcement procedure in the BBR's design, the voice and accountability measure, the Dummy variable reflecting if a country is a federal country, the dummy reflecting if a country is member of a currency union, the number of years the chief executive hold, a proxy for voter preferences, the first lag of the interest payments on debt (expressed in percent of total public expense), the first lag of public balance (expressed in percent of GDP). The sign reported next to the identified determinants of BBR indicates whether the factor affects positively or negatively BBR's compliance.

Dummy variable for Crisis has a negative effect on BBR's compliance. It suggests that it is difficult for governments to comply with fiscal rules during worst economic periods. The presence of escape clauses makes governments tempted to not comply BBR. Governments seem tempted to relax because of the presence of these escape clauses. It constitutes an empirical evidence that escape clauses drive government behavior and thus matter in the choice of compliance definition. On the contrary, the presence of sanctions for non-compliance positively affects BBR's compliance. It means that governments seem to consider with caution the potential application of financial sanctions if they deviate from their objective. The lagged value of interest payments on debt increases the compliance in the next year, suggesting that governments try to implement effort to comply to send a positive signal to financial market. Without surprise, the lagged value of public balance positively affects the BBR's compliance because it is easier to comply fiscal rule when public finance is in good health. Finally, one of our two proxies for voter preferences appears significant. We tested two proxies of voter preferences coming from our latent factor computation. The significance of one these two factors reflecting voter preferences, suggests that we must take into account voter preferences when we assess fiscal rules effects. Indeed, the voter preferences seem to increase the BBR's compliance, reflecting an average preference of the voters for disciplined governments. The number of years of a chief executive has been in place is positively linked with BBR's compliance. If voters indeed prefer complier-government, a disciplined chief executive will stay longer and increase BBR's compliance. Otherwise, the significance but negative effect of the "voice and accountability" may reflect that the place offered by more democratic economies to public policies debates may sometimes act as a brake on fiscal discipline.

Table 3.2: **Compliance determinants**

| LASSO and BOOSTING common determinants |
|---|
| Dummy variable for crisis (-) |
| Dummy variable for Well-specified escape clause (-) |
| Dummy variable for Formal enforcement procedure (+) |
| Voice and Accountability (-) |
| Dummy variable for Federal country (+) |
| Dummy variable for member of a currency union (+) |
| Years chief executive (+) |
| The First proxy for voter preferences (+) |
| $lag-1$ interest payments (in % of expense) (+) |
| $lag-1$ of Public Balance (in % of GDP) (+) |

Note : Years chief executive reflects the number of years the chief executive was in office . Election system takes value 2 for parliamentary system, 1 for Assembly-elected President and 0 for Presidential system (see Database of Political Institutions 2015 (2016) for further details). Only the ten common indicators are reported: L2-Boosting retained 10 key determinants and Lasso retained 15 (among these fifteen key determinants ten are the same as in L2-Boosting) . The signs (+) and (-) reflects the impact sign of the variable on BBR-compliance.

### 3.5.2 Average Treatment Effect on social welfare channels

Table 3.3 presents the Average Treatment Effect (ATE) of BBR's compliance on our variables of interest. We decompose our results in a first part that summarizes the ATE on the macroeconomic variables while the second part reveals the ATE on government performance and inequalities indicators. All our results are stable across feature selections approaches used in the first step of our DML algorithm. Nevertheless, the RMSE for the dependent variables provided by L2-Boosting is lowest in every case, showing that it is the best model.

The Table 3.3 -part 1- highlights that, according to literature which links fiscal rules and fiscal discipline [22], the BBR's compliance increases *on average* the general government public balance by 0.5 percentage points (hereafter pp) (column 1). Nevertheless, BBR compliers seem to not benefit from lower interest rate on public debt since the corresponding ATE is not significant in column 2. This suggests that compliance does not send a positive signal-effect to financial markets. However, the Chapter 1 showed that the fiscal rules presence reduces the interest rate on debt. Finally, the simple presence of fiscal rules matters as a signal effect for financial markets, but fiscal rule compliance

[22] See Section 3.1 and 3.2 for discussion

does not imply any difference. This finding highlights that the definition of fiscal rules performance retained may drive the conclusions.

The total public spending decrease by 0.125 pp for BBR compliers while general government investment (Gross Fixed Capital Formation (GFCF)) increases by 0.263 pp *on average* as showed by, respectively, significant and negative ATE (column 4) for total public spending and significant and positive ATE (column 3) for general government GFCF. As explanation, governments operate a cut in government final consumption to promote BBR's compliance as we can see a negative and significant ATE on general government final consumption in column 5. The final negative effect on public spending is the result of a cut in public consumption of fixed capital.

Through the increase in GFCF, compliers seem to expect economic growth benefits. They indeed present a GDP growth expectation 0.6 pp higher than for non-compliers, as suggested in column 6. However, in practice, their spending re-allocation do not provide higher GDP growth in the next year as suggested by column 7 where BBR's compliance has no impact on future GDP.

Table 3.3 -part 2- reports that the BBR's compliance has no effect on Government Effectiveness and Government Efficiency. We expected that fiscal rules' compliance forces government to spend in a better way, taking care of each unit of money spent and thus increase government efficiency. We also expected that government favor spending performance in order to insure favorable economic conditions and thus increase government effectiveness. Nevertheless, we observe that ATE associated with both government effectiveness and government efficiency are not significant. Governments reduce social spending but increase GFCF at the same time; two actions going on the opposite side that finally lead to a zero-effect on the government performance. A major result is found in column 3 of Table 3.3 part 2: we observe a positive and significant ATE on Gini index. Since Gini index is, by definition, an index between 0 and 1 without common units, it couldn't be interpreted as variables expressed in percent of GDP. The BBR's compliance leads to an increase around 0.09 units in the Gini index. By forcing compliance, but by simultaneously trying to increase public GFCF, government go beyond the trade-off between BBR's compliance and growth objectives and conduct to a side-effect on social spending. Some social spending is included in the government final consumption expenditure which is reduced by the BBR's compliance. We thus observe an increase in inequalities measured through the Gini index. 0.09 unit of Gini index represents 9% of the index values' range.

In that sense, compliance may explain around 9% of the differences in Gini index between compliers and non-compliers. As suggested by the last column of Table 3.3 part 2, the poorest are affected by the spending re-allocation. Finally, the side-effects observed in public expenditure impact both inequalities and poverty, suggesting that government may to face a trade-off between fiscal rules' compliance and social objectives.

Table 3.4 shows the robustness test by removing observations for the UK and Hungary on which we set hypotheses in Section 3.3.1. Our results still hold with the two methods, and L2-boosting is still being the best model regarding the RMSE measure.

Table 3.5 provides results removing observations-years where an escape clause holds. We see that all results are still the same except for the GDP per capital growth in $t+1$. A more flexible definition of fiscal rules' compliance, allowing escape clause to matter, is favorable for economic growth. Consequently, escape clauses matter for compliance definition in two dimensions: i) escape clauses affect compliance itself by promoting compliance (according to results in Section 3.5.1); ii) escape clauses affect BBR's effect since if we allow flexibility in compliance definition, GDP growth appears higher.

Nevertheless, adopting a flexible compliance's definition does not change the adverse effects on inequality as we do not observe any improvement in government performance. Unsurprisingly, as public spending is not better allocated, there is no improvement in social conditions. So we cannot say that complying with the BBR would be beneficial for economic welfare because: i) a sophisticated and flexible definition of compliance is necessary to find a positive effect (only) on economic growth, ii) this flexibility in compliance's definition is not sufficient as inequality is always increased by compliance. Governments seeking to ensure fiscal discipline must therefore carefully consider the importance of the social dimension before redirecting their spending. Finally, BBR's compliance may not damageable for the economic area but for the social area. Such result reinforces our highlight suggesting that the side-effect on public spending composition is negative for social welfare and governments seem not to face a "Compliance vs GDP growth trade-off" but they deal with a "Compliance vs Social objectives trade-off".

Table 3.3: **ATE of Budget Balance Compliance with 5-fold cross-validation**

Part 1: ATE on Macroeconomic channels

| DML Estimator \ Dependent Variable | GG Public Balance | Interest payments (% of expense) | GG GFCF (in % of GDP) | Total spending (in % of GDP) | GG final consumption (in % of GDP) | GDP per cap. expectation | GDP per cap. Growth in $t+1$ |
|---|---|---|---|---|---|---|---|
| LASSO | 0.534*** | 0.058 | 0.263*** | -0.125*** | -0.107*** | 0.601*** | 0.140 |
| | (0.100) | (0.049) | (0.077) | (0.034) | (0.028) | (0.170) | (0.098) |
| RMSE y | 0.532 | 0.338 | 0.370 | 0.172 | 0.202 | 0.402 | 0.557 |
| BOOSTING | 0.481*** | 0.108 | 0.266*** | -0.095*** | -0.141*** | 0.526*** | 0.077 |
| | (0.087) | (0.030) | (0.068) | (0.023) | (0.029) | (0.151) | (0.109) |
| RMSE y | 0.392 | 0.234 | 0.283 | 0.125 | 0.136 | 0.341 | 0.403 |

Part 2: ATE on government performance and inequalities channels

| DML Estimator \ Dependent Variable | Government Effectiveness | Government Efficiency | Gini Index | Poverty headcount ratio at 1,90$ a day (2011 PPP) (% of population) |
|---|---|---|---|---|
| LASSO | -0.014 | 0.128 | 0.087* | 0.079** |
| | (0.033) | (0.140) | (0.072) | (0.035) |
| RMSE y | 0.147 | 0.635 | 0.344 | 0.216 |
| BOOSTING | -0.019 | 0.099 | 0.032* | 0.049** |
| | (0.031) | (0.133) | (0.065) | (0.036) |
| RMSE y | 0.118 | 0.284 | 0.274 | 0.192 |

Note: "RMSE y" refers to Root Mean Square Error of the dependent variable. GG = General Government, GFCF = Gross Fixed Capital Formation. The median standard error across the splits is reported in brackets.

Source: Author.

Table 3.4: **Robustness ATE of Budget Balance Compliance with 5-fold cross-validation: without observations related to hypotheses set by the author in 3.3.1**

Part 1: ATE on Macroeconomic channels

| DML Estimator \ Dependent Variable | GG Public Balance | Interest payments (% of expense) | GG GFCF (in % of GDP) | Total spending (in % of GDP) | GG final consumption (in % of GDP) | GDP per cap. expectation | GDP per cap. Growth in $t+1$ |
|---|---|---|---|---|---|---|---|
| LASSO | 0.470*** | 0.020 | 0.231*** | -0.107*** | -0.172*** | 0.580*** | 0.120 |
| | (0.096) | (0.052) | (0.068) | (0.029) | (0.037) | (0.160) | (0.127) |
| RMSE y | 0.510 | 0.359 | 0.348 | 0.156 | 0.195 | 0.385 | 0.560 |
| BOOSTING | 0.452*** | 0.072 | 0.257*** | -0.095*** | -0.123*** | 0.581*** | 0.039 |
| | (0.079) | (0.025) | (0.063) | (0.022) | (0.032) | (0.132) | (0.090) |
| RMSE y | 0.387 | 0.248 | 0.281 | 0.125 | 0.150 | 0.329 | 0.400 |

Part 2: ATE on government performance and inequalities channels

| DML Estimator \ Dependent Variable | Government Effectiveness | Government Efficiency | Gini Index | Poverty headcount ratio at 1,90$ a day (2011 PPP) (% of population) |
|---|---|---|---|---|
| LASSO | -0.0005 | 0.125 | 0.079* | 0.087** |
| | (0.032) | (0.135) | (0.071) | (0.036) |
| RMSE y | 0.153 | 0.661 | 0.359 | 0.200 |
| BOOSTING | 0.002 | 0.064 | 0.058* | 0.031** |
| | (0.029) | (0.146) | (0.069) | (0.034) |
| RMSE y | 0.121 | 0.329 | 0.298 | 0.197 |

Note: "RMSE y" refers to Root Mean Square Error of the dependent variable. GG = General Government, GFCF = Gross Fixed Capital Formation. The median standard error across the splits is reported in brackets.

Source: Author.

Table 3.5: **Robustness ATE of Budget Balance Compliance with 5-fold cross-validation: without observations that did not comply with their BBR but escape clauses existed**

Part 1: ATE on Macroeconomic channels

| DML Estimator \ Dependent Variable | GG Public Balance | Interest payments (% of expense) | GG GFCF (in % of GDP) | Total spending (in % of GDP) | GG final consumption (in % of GDP) | GDP per cap. expectation | GDP per cap. Growth in $t+1$ |
|---|---|---|---|---|---|---|---|
| LASSO | 0.431*** | 0.084 | 0.242*** | -0.088*** | -0.117*** | 0.551*** | 0.247*** |
| | (0.091) | (0.040) | (0.067) | (0.025) | (0.032) | (0.152) | (0.084) |
| RMSE y | 0.509 | 0.359 | 0.379 | 0.151 | 0.183 | 0.436 | 0.436 |
| BOOSTING | 0.514*** | 0.107 | 0.237*** | -0.099*** | -0.139*** | 0.527*** | 0.156*** |
| | (0.084) | (0.041) | (0.071) | (0.023) | (0.037) | (0.151) | (0.094) |
| RMSE y | 0.387 | 0.246 | 0.286 | 0.127 | 0.157 | 0.317 | 0.392 |

Part 2: ATE on government performance and inequalities channels

| DML Estimator \ Dependent Variable | Government Effectiveness | Government Efficiency | Gini Index | Poverty headcount ratio at 1,90$ a day (2011 PPP) (% of population) |
|---|---|---|---|---|
| LASSO | -0.0005 | 0.125 | 0.079* | 0.087** |
| | (0.032) | (0.135) | (0.071) | (0.036) |
| RMSE y | 0.153 | 0.661 | 0.359 | 0.200 |
| BOOSTING | -0.013 | 0.027 | 0.028* | 0.064* |
| | (0.031) | (0.175) | (0.088) | (0.045) |
| RMSE y | 0.120 | 0.557 | 0.312 | 0.177 |

Note: "RMSE y" refers to Root Mean Square Error of the dependent variable. GG = General Government, GFCF = Gross Fixed Capital Formation. The median standard error across the splits is reported in brackets.
Source: Author.

### 3.5.3 Technical and policy recommendations

Following our main results, the first implication for future research is technical. Indeed, the significance of voter preferences proxy suggests that taking care of voter preferences when assessing fiscal rules performance highly matter. Neglecting this variable could lead to an omission bias, which appears important for all models estimating average treatment effects. As developed in Chapter 1, the accurate specification of the treatment itself is key to achieve the conditional independence assumption described by Rosenbaum and Rubin [1983]. Statistical robustness tests may be insufficient to cover such omission bias when assessing fiscal rules performance, as discussed in Wyplosz [2012] and Heinemann et al. [2018]. Political database such as the Database on Political Institution (DPI, Cruz et al. [2020]) should be seriously considered and used to construct variables to proxy and control voter preferences.

The other technical recommendation is to extend the use of models that account for reverse causality and omission bias at the same time. The use of causal Machine Learning to estimate inference parameter may offer opportunities for future research. Even tough, Machine Learning is mostly famous for forecasting and classification, it should also be considered as an alternative for econometrics in causal estimation. The Double/Debiased Machine Learning model developed by Chernozhukov et al. [2018] that we used in this chapter presents several advantages such as testing for a larger number of predictors than standard econometrics approach, and it thus reduces the potential omission bias. The orthogonalization procedure and the use of lagged macroeconomic variables in our model discard reverse causality bias. The risk of overfitting is avoided by cross-validation procedure. Consequently, the use of the combination of these techniques proposes a Machine Learning tool as a solution with strong asymptotic properties for causal estimation. We thus support the use of such approaches for future studies in fiscal issues as well as in other macroeconomic topics.

On the other hand, we must put some words on policy recommendations. Due to the importance of voter preferences and the number of years that a chief executive stays on office, as fiscal rules compliance determinants, governments should consider with caution the importance that fiscal discipline represents for elective purposes. We first expected that governments may be tempted to run deficit to increase public spending to carry favor from electors in order to be re-elected. Nevertheless, when electors prefer disciplined governments such behavior no longer holds. This appears even more relevant because modern societies are increasingly informed and not easily fooled by government elective strategies.

Our findings regarding the side-effects of fiscal rules compliance should serve as a warning as well as a guide for the future design of fiscal rules. The flexibility of fiscal rules compliance definition (by considering escape clauses) seems to limit the negative effects we found by improving the GDP growth. However, it is not sufficient to limit the negative impact on social expenditure (in particular on social transfers). The coming years will not be able to ignore the amount of public debt accumulated that followed covid-19 crisis. Fiscal rules will therefore have an important role to play in restoring fiscal discipline. But, this cannot be done without serious considerations of social spending and inequalities, as the pandemic crisis has also increased inequalities by affecting more some sectors than others[23]. Thus, the future of fiscal rules must be achieved through thoughtful and discussed reforms, favoring fiscal discipline while preserving productive spending (investment) without damaging social spending. There is no miracle solution, but improvements are possible. In particular, the multiplicity and complexity of fiscal rules, as in the Stability and Growth Pact in the EU, may make fiscal rules inefficient but also not credible. A simplification of fiscal framework using more flexible rules regarding the definition of their target indicators could be considered. We are thinking, in particular, of over-the-cycle rules or a Golden rule. Such fiscal rules that are more flexible by definition, could include sanctions in case of deviation from the rule. Indeed, a country that does not comply with a flexible rule could turn out to be far too lax in relation to the margins already authorized by the rule.

## 3.6 Conclusion

The chapter provides an assessment of national Budget Balance Rules compliance side-effect on social welfare channels indicators. It uses the Double/Debiased Machine Learning methodology including LASSO or Boosting feature selection algorithms as robustness test. All the results do not depend on the shrinking algorithm choice since results are consistent across feature selection estimators. From the feature selection step, a set of

[23]For example, restaurants and shops have been on partial unemployment for a very long time, while other jobs have been able to telework without loss of pay.

key determinants for BBR's compliance is identified including voter preferences (suggesting that voter preferences need to be taking into account in fiscal rules analyses). Such empirical results suggest that Wyplosz [2012]'s bias matters.

Finally, average treatment effect results, from the second step, showed that governments with national Budget Balance Rules seem to try to overcome the trade-off between BBR's compliance and Growth objectives. Governments conduct public investment and achieve BBR's compliance at the same time. Instead of an arbitration between compliance and economic growth, governments operate a re-allocation of spending. Governments seem to favor Gross Fixed Capital Formation but decrease government Final Consumption that includes social spending. Consequently, BBR's compliance seem to have an increasing effect on inequalities and this effect affect more the poorest classes as suggested by the impact on the poverty head account ratio. Finally, empirical findings provide of side-effects of fiscal rules strict compliance. Nevertheless, by relaxing the compliance definition, we finally found similar conclusion as in Blundell et al. [2011], that fiscal rules may support economic growth. The side-effects of fiscal rules' compliance operate through public spending composition by decreasing social spending. Consequently, we should not recommend abandoning fiscal rules and their rigorous application but to better design them. Flexible fiscal rules have been largely discussed in the literature (see Eyraud et al. [2018], Caselli et al. [2018]) and they may be a solution to limit fiscal rules' compliance side-effect. Indeed, the inclusion of escape clauses may have positive effects on economic growth, but it does not appear sufficient to limit side-effect on inequalities. But, Debrun and Jonung [2019] proposed a fiscal-Taylor rule following an over-cycle expenditure benchmark, while others as Creel et al. [2014] argue in favor of the Golden rule. Both seem to work against the weakness regarding public social spending but the fiscal rules should be precisely defined, including a social related objective. However, an expenditure benchmark or a Golden rule require a harmonization of governments accounting, especially for the members of a common currency union as the euro area. This leads to a higher debate on what should be considered as a productive expenditure and how to compute government consumption of fixed capital (see Schreyer [2003] for discussion on productive capital and countries computational hypotheses).

Chapter 2 already supported the use of Machine Learning but for forecasting purposes. The results of Chapter 3 also launch the discussion on the use of Machine Learning in the econometric field (Athey [2018]). Indeed, this chapter proposed a robusr causal Machine Learning estimator against standard econometrics biases such as reverse causality or omission bias. Consequently, Machine Learning may be seriously considered as a useful tool in causal inference economic studies.

# Appendices

**Appendix 1: Source of Budget Balance Rules' targeted values**

| Country | Years | Source for Budget Balance Rule's Target |
|---|---|---|
| Chile | 2004-2015 | IMF World Economic Outlook Database 2018 |
| Costa-Rica | 2004-2015 | Fiscal balance comes from World Bank except in 2015 where Fiscal Balance comes from Banco Central de Costa Rica (BCCR) and Gross Fixed Capital Formation comes from IMF Investment and Capital Stock dataset 1960-2015 |
| Denmark | 2004-2015 | IMF World Economic Outlook Database 2018 |
| Estonia | 2004-2015 | IMF World Economic Outlook Database 2018 |
| Finland | 2004-2015 | Eurostat |
| Germany | 2004-2010 | Eurostat |
| Germany | 2011-2015 | IMF World Economic Outlook Database 2018 |
| Hungary | 2004-2015 | IMF World Economic Outlook Database 2018 |
| Indonesia | 2004-2015 | IMF World Economic Outlook Database 2018 |
| Israel | 2004-2015 | IMF World Economic Outlook Database 2018 |
| Japan | 2004-2015 | IMF World Economic Outlook Database 2018 |
| Malaysia | 2004-2015 | IMF World Economic Outlook Database 2018 and Gross Fixed Capital Formation comes from IMF Investment and Capital Stock dataset 1960-2015 |
| New Zealand | 2004-2015 | New Zeland Treasury *Fiscal Time Series Historical Indicators 1972 - 2018* |
| Peru | 2004-2015 | IMF (Peru: Selected Issues Paper, IMF, 2012, number 12-27) and Banco Central de Reserva del Peru (BCRP) |
| Spain | 2004-2015 | IMF World Economic Outlook Database 2018 |
| Sweden | 2004-2015 | IMF World Economic Outlook Database 2018 |
| Switzerland | 2004-2015 | IMF World Economic Outlook Database 2018 |
| United Kingdom | 2004-2009 | Eurostat |
| United Kingdom | 2010-2015 | IMF World Economic Outlook Database 2018 |

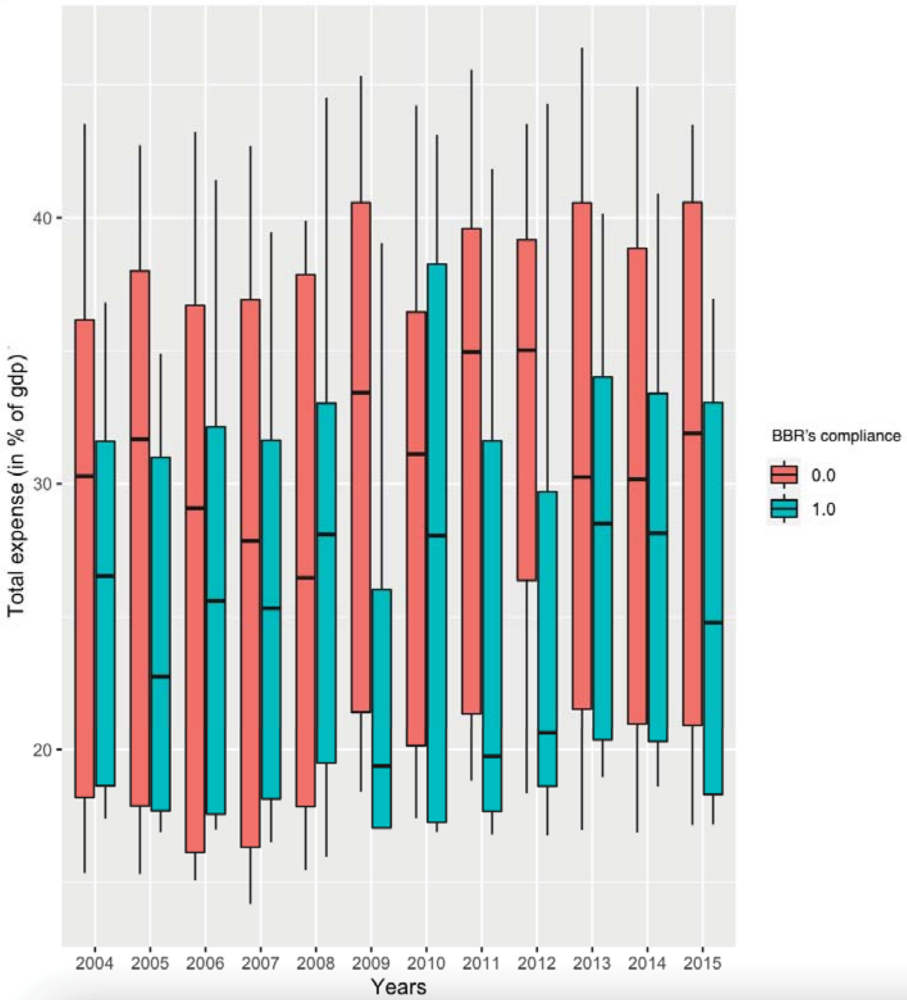
Source: Author.

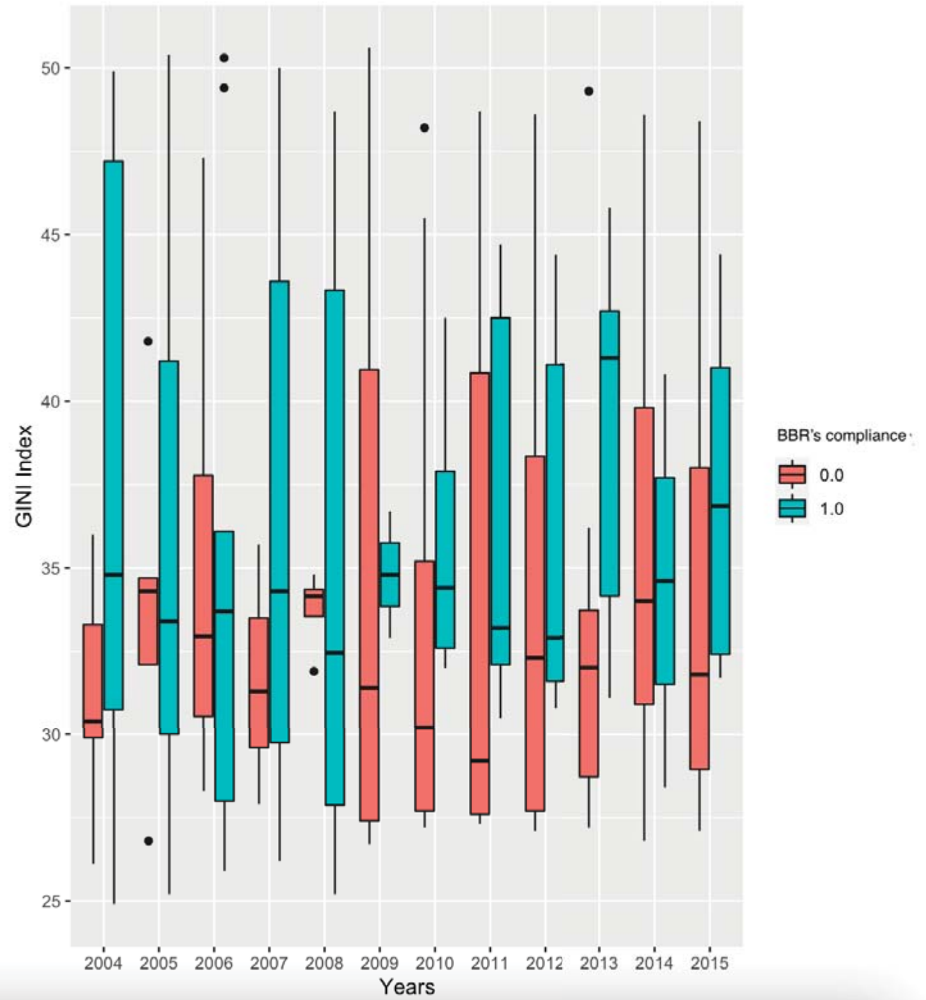
| Country | Constrained variable | Target Value | Period | Comments | Level of Government constrained |
|---|---|---|---|---|---|
| Chile | Structural Balance | 1 | 2001-2007 | | Central |
| | Structural Balance | [0.5; -2] | 2008-2015 | Rule defined in a range | Central |
| Costa Rica | Budget Balance excluding gross investment | 0 | 2001-2015 | Golden Rule | Central |
| Denmark | Structural Balance | 0.5 | 2001-2011 | | General |
| | Structural Balance | -0.5 | 2014<br>2015 | | General |
| Estonia | Structural Balance | 0 | 1993<br>2011 | | General |
| | Structural Balance | 0 | 2012 | | General |
| Finland | Structural Balance | [0;1] | 1999-2013 | 1% between 2007 and 2011 | Central |
| | Budget Balance (Total) | [-2.75; -2.5] | 1999-2008 | -2.75% between 1999 and 2002 | Central |
| | Budget Balance (Total) | -1 | 2011 | | Central |
| Germany | Budget Balance excluding net investment | 0 | 1969-2010 | | Central |
| | Structural Balance | -0.35 | 2011 | | Central |
| Hungary | Primary Balance | 0 | 2004<br>2009 | | General |
| | Annual changes of Primary Balance | 0 | 2010 and 2011 | | General |
| | Primary Balance | 0 | 2009<br>2011 | Not included in our analysis | Central |
| | Structural deficit above 0.5% (because debt is higher that 60% as described in TSCG) | -0.5 | 2012-2015 | Transpose in national law from TSCG, interpreted as national BBR by hypothesis here | General |
| Indonesia | Budget Balance (Total) | -3 | 1967-2015 | | General |
| Israel | | | | Not included due to annual change in the targeted value. Not a numerical rule. | |
| Japan | Budget Balance excluding net investment | 0 | 1990--2015 | Golden Rule | Central |
| Malaysia | Budget Balance excluding net investment | 0 | 1959-2015 | Golden Rule | Central |
| New Zealand | Budget Balance excluding net investment | 0 | 1994-2015 | Golden Rule | General |
| Peru | Budget Balance | [-1; 2] | 2000<br>2013 | | Central |
| | Structural Balance | -1 | 2014 | | General |
| Spain | Budget Balance (Total) | [-2;0] | 2003<br>2011 | Limit related to GDP growth | General |
| Sweden | Budget Balance | [1;2] | 2000 | Only 1% since 2007 | General |
| Switzerland | Structural Balance | 0 | 2003 | | Central |
| United Kingdom | Budget Balance excluding net investment | 0 | 1997<br>2008 | Golden Rule | General |
| | Annual changes in Budget Balance (Total) | 0 | 2010 | | General |

Note: BBR = Budget Balance Rule. We stop all reported periods in 2015 because IMF Fiscal Rules Database only reports fiscal rules until 2015. It does not mean that fiscal rules are no more in force after 2015. Source: Caselli et al. [2018], Reuter [2019], Eyraud et al. [2018], but we assume some differences for Hungary, Japan and United Kingdom developed in section 3.1 and robustness tests are implemented in section 3.5.

Source: Author.

**Appendix 2. Fiscal rules included in our analysis between 2004 and 2015**

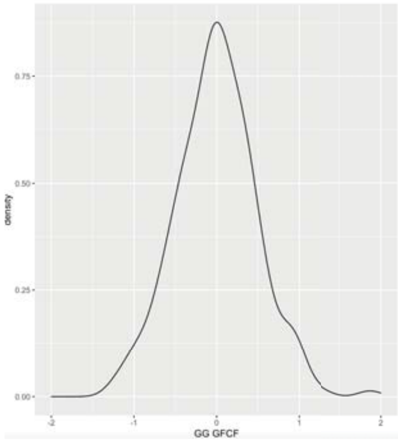



Note: BBR = Budget Balance Rule. “0” means BBR non-compliance and “1” means BBR’s compliance.

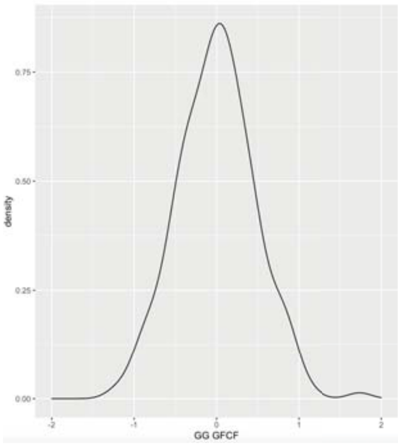
Source: Author.

**Appendix 3. Comparison of Public Spending and Gini index between BBR compliers and BBR non-compliers by quantiles**




Note: GFCF stands for Gross Fixed Capital Formation. All data are standardized before applying ML algorithm.
Source: Author.

**Appendix 4. Distribution of General Government GFCF fitted values resulting from LASSO feature selection**




Note: GFCF stands for Gross Fixed Capital Formation. All data are standardized before applying ML algorithm.
Source: Author.

**Appendix 5. Distribution of General Government GFCF fitted values resulting from Boosting feature selection**

# Bibliography

D. Acemoglu. Why not a political coase theorem? Social conflict, commitment and politics. *Journal of Comparative Economic Studies*, pages 620–652, 2003. 128

A. Afonso, L. Schuknecht, and V. Tanzi. Public Sector Efficiency: Evidence from new EU member states and Emerging market. *ECB Working Paper No. 581*, 2006. 136

A. Afonso, J. T. Jalles, and A. Venâncio. Taxation and Public Spending Efficiency: An International Comparison. *Working Papers REM 2019/80*, 2019. 136

S. R. Aiyagari and E. R. McGrattan. The optimum quantity of debt. *Journal of Monetary Economics*, 1998. 129

K. Arrow. Social Choice and Individual Values. *Yale University Press*, 1951. 128

S. Athey. The Impact of Machine Learning on Economics. *Chapter in NBER book The Economics of Artificial Intelligence: An Agenda*, pages 507–547, 2018. 157

S. Athey and G.W. Imbens. The state of applied econometrics: Causality and policy evaluation. *The Journal of Economic Perspectives*, 2017. 143

S. Athey, G. Imbens, and S. Wager. Approximate residual balancing: De-biased inference of average treatment effects in high-dimensions. *arXiv:1604.07125v3 . arXiv*, 2016. 143

S. Athey, J. Tibshirani, and S. Wager. Generalized random forests. *arXiv:1610.01271*, 2017. 143

R. Beetsma, X. Debrun, X. Fang, Y. Kim, V. Lledó, S. Mbaye, and X. Zhang. Independent Fiscal Councils: Recent Trends and Performance. *IMF Working Paper 18/68*, 2018. 131

A. Belloni, V. Chernozhukov, D. Chetverikov, C. Hansen, and K. Kato. High-Dimensional Econometrics and Regularized GMM. *Cornelle University, arXiv:1806.01888v2*, 2018. 126

M. Bergman, M.M. Hutchison, and S.E. Hougaard Jensen. Promoting sustainable public finances in the European Union: The role of fiscal rules and government efficiency. *European Journal of Political Economy*, 44:1–19, 2016. 125

R. Blundell, M. Costa Dias, R. Joyce, and X. Xu. The impact of the European Union fiscal rules on economic growth. *Journal of Macroeconomics*, pages 313–326, 2011. 129, 157

R. Blundell, M. Costa Dias, R. Joyce, and X. Xu. COVID-19 and Inequalities. *Fiscal Studies*, pages 291–319, 2020. 125

H. Bohn and R.P. Inman. Balanced Budget Rules and Public Deficits: Evidence from the U.S. States. NBER Working Paper No. 5533, 1996. 129

J. B. Burbidge. Government Debt in an Overlapping-Generations Model with Bequests and Gifts. *The American Economic Review*, pages 222–227, 1983. 129

P. Bühlmann and B. Yu. Boosting With the L2 Loss: Regression and Classification. *Journal of the American Statistical Association*, 2003. 147

F. Caselli, L. Eyraud, A. Hodge, F. Diaz Kalan, Y. Kim, V. Lledó, S. Mbaye, A. Popescu, W. H. Reuter, J. Reynaud, E. Ture, and P. Wingender. Second-Generation Fiscal Rules : Balancing Simplicity, Flexibility, and Enforceability - Technical Background Papers . *Staff Discussion Notes No. 18/04*, 2018. 131, 134, 157, 160

V. Chernozhukov, D. Chetverikov M. Demirer, E. Duflo, C. Hansen, and W. Newey. Double/Debiased/Neyman Machine Learning of Treatment Effects. *American Economic Review*, 2017. 126, 143, 144, 145

V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, Ch. Hansen, W. Newey, and J. Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 2018. 126, 143, 145, 155

J-L. Combes, A. Minea., P.N. Sawadogo, and C. Vinturis. Can fiscal rules curb income inequality? Evidence from developing countries. *CERDI Etudes et Documents No. 25*, 2019. 130, 132

J-L. Combes, X. Debrun, A. Minea, and R. Tapsoba. Inflation Targeting, Fiscal Rules and the Policy Mix: Cross-effects and Interactions. *The Economic Journal*, 128:2755–2784, 2018. 125

J. Creel, P. Veroni, and F. Saraceno. Has the Golden Rule of Public Finance Made a difference in the UK. *hal-00972843*, 2014. 157

C. Cruz, P. Keefer, and C. Scartascini. Database of Political Institutions. *Inter-American Development Bank, Department of Research and Chief Economist*, 2020. 155

X. Debrun and L. Jonung. Under threat: Rules-based fiscal policy and how to preserve it. *European Journal of Political Economy*, 57:142–157, 2019. 157

X. Debrun and M.S. Kumar. The Discipline-Enhancing Role of Fiscal Institutions: Theory and Empirical Evidence. *IMF Working Paper No. 07/171*, 2007. 137

X. Debrun, L. Moulin, A. Turrini, J. Ayuso i casals, and M. Kumar. Tied to the Mast? National Fiscal Rules in the European Union. *Economic Policy*, 23:297–362, 2008. 125, 133

M. Delgado-Téllez, V. D. Lledó, and J. J. Pérez. On the Determinants of Fiscal Non-Compliance: An Empirical Analysis of Spain's Regions. *IMF Working Paper 17/5*, 2017. 125, 136

G. Everaert, F. Heylen, and R. Schoonackers. Fiscal policy and tfp in the oecd: measuring direct and indirect effects. *Empirical Economics*, pages 605–640, 2015. 135

L. Eyraud, X. Debrun, A. Hodge, V. Lledó, and C. A Pattillo. Second-Generation Fiscal Rules : Balancing Simplicity, Flexibility, and Enforceability. *Staff Discussion Notes No. 18/04*, 2018. 123, 131, 157, 160

M. Fajeau. Too much finance or too many weak instruments? *International Economics*, 2021. 126

M. Flodén. The effectiveness of government debt and transfers as insurance. *Journal of Monetary Economics*, 2001. 129

J.H. Friedman, T. Hastie, and R. Tibshirani. Additive Logistic Re- gression: A Statistical View of Boosting. *The Annals of Statistics*, 2000. 147

R. Frisch and F.V. Waugh. Partial time regression as compared with individual trends. *Econometrica*, pages 387–401, 1933. 144

P. Funk and C. Gathmann. Voter preferences, direct democracy and government spending. *European Journal of Political Economy*, 2013. 137

P. Gogas, T. Papadimitriou, and A. Agrapetidou. Forecasting bank failures and stress testing: A machine learning approach. *International Journal of Forecasting*, 34:440–455, 2018. 144

S. Gosh and U. Roy. Fiscal policy, long-run growth, and welfare in a stock-flow model of public goods. *Canadian Journal of Economics/Revue canadienne d'économique*, pages 742–756, 2004. 128

M. Guerguil, P. Mandon, and R. Tapsoba. Flexible fiscal rules and countercyclical fiscal policy. *Journal of Macroeconomics*, 52:189–220, 2017. 129

J. Hartwig and J.E. Strum. Do fiscal rules breed inequality? First evidence for the EU. *Economic Bulletin*, 2019. 130, 132

W. Hediger. Sustainable development and Social Welfare. *Ecological Economics*, pages 481–492, 2000. 128

F. Heinemann, M-D. Moessinger, and M. Yeter. Do fiscal rules constrain fiscal policy? A meta-regression-analysis. *European Journal of Political Economy*, 51:69–92, 2018. 126, 137, 155

W. Härdle, Y-J. Lee, D. Schäfer, and Y-R. Yeh. Variable Selection and Oversampling in the use of Smooth Support Vector Machines for Predicting the Default Risk of Companies. *Journal of Forecasting*, 28:512–534, 2009. 144

G. Kopits and S.A. Symansky. Fiscal Policy Rules . *Occasional Paper No. 162*, 1998. 131

S. Kuznet. Economic Growth and Income Inequality. *The American Economic Review*, pages 1–28, 1955. 130

M. Larch and S. Santacroce. Numerical compliance with EU fiscal rules: The compliance database of the Secretariat of the European Fiscal Board (Database). 2020. 125

M. Larch, E. Orseau, and W. van der Wielen. Do EU Fiscal Rules Support or Hinder Counter-Cyclical Fiscal Policy?). *Journal of International Money and Finance*, 2021. 129, 132, 136

P. Lawler. Monetary policy, central bank objectives, and social welfare with strategic wage setting. *Oxford Economic Papers*, page 94–113, 2001. 128

M.C. Lovell. Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association*, pages 933–1010, 1963. 144

F. Marini. Banks, financial markets, and Social Welfare. *Journal of Banking and finance*, pages 2557–2575, 2005. 128

A. H. Maslow. Motivation and personality. *(2nd ed.). New York: Harper and Row.*, 1970. 124

J. Midgley. Growth, Redistribution, and Welfare: Toward Social Investment. *University of Chicago Press*, 1999. 128, 129

S. Mullainathan and J. Spiess. Machine learning: an applied econometric approach. *The Journal of Economic Perspectives*, 2017. 143

W.H. Reuter. National numerical fiscal rules: not complied but still effective? *European Journal of Economy*, 39:67–81, 2015. 131

W.H. Reuter. When and why do countries break their national fiscal rules? *European Journal of Political Economy*, 57:125–141, 2019. 125, 134, 136, 160

P.R. Rosenbaum and D. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 1983. 127, 144, 155

A. Sacchi and S. Salotti. The impact of national fiscal rules on the stabilisation function of fiscal policy. *European Journal of Political Economy*, 37:1–20, 2015. 129

P. Samuelson. Welfare Economics and International Trade. *The American Economic Review*, pages 261–266, 1938. 128

A. Schaechter, T. Kinda, and N. Budinaand A. Weber. Fiscal Rules in Response to the Crisis—Towards the 'Next-Generation' Rules. A New Dataset. *IMF Fiscal Affairs Department*, 2012. 133, 134

P. Schreyer. CAPITAL STOCKS, CAPITAL SERVICES AND MULTI-FACTOR PRODUCTIVITY MEASURES. *OECD Economic Studies, No. 37, 2003/2*, 2003. 157

D. M. Smith. Who gets what Where, and How: A welfare focus for human geography. *Geography*, pages 289–297, 1974. 128

R. Tapsoba. Do National Numerical Fiscal Rules really shape fiscal behaviours in developing countries? A treatment effect evaluation. *Economic Modelling*, 29:1356–1369, 2012. 125

H.R. Varian. Big data: New tricks for econometrics. *The Journal of Economic Perpectives*, 2014. 143

C. Wyplosz. Fiscal Rules: Theorical issues and historical experience. *NBER Working Paper No. 17884*, 2012. 127, 129, 137, 155, 157

# Chapter 4

# Statistical insights behind fiscal rules performance assessment

"(One existing) theory advocates the implementation of a "Golden rule of public finances" in order to reduce the governments' bias for running excessive deficits: current expenditure must be financed through taxatons, while investment which will benefit future generations may be financed through borrowing. It is however difficult to measure investment. " Mathieu and Sterdyniak [2013]

## 4.1 Introduction

The debate around the future of fiscal rules is a burning discussion, especially in the EU context where it concerns serious proposals to reform the Stability and Growth Pact (SGP). These considerations require political agreements as well as efforts in the harmonisation and the transparency of the statistical methods used by countries regarding the indicators targeted by fiscal rules. This Chapter 4 provides statistical highlights and recommendations for the future of fiscal rules as if one envisages a future where fiscal rules are still present, it needs to take care of their design and the accuracy in the series of their targeted indicators to avoid any misinterpretation or error in fiscal rules performance assessment. This chapter is interested in the impact of statistical methodologies of countries national account on the fiscal rules performance assessment. In particular, it addresses the following question: "How public net investment measurement methodologies may affect the fiscal rules performance assessment?".

To answer this problem, we will focus on the sensitivity of the Golden rule of public finance to the statistical assumptions behind the measurement of general government consumption of fixed capital (thereafter CFC) which enters in the definition of the Golden rule target measure. This chapter seems more relevant under the prism of the current pandemic crisis, which followed the Sovereign Debt Crisis (2010) and the GFC (2008), that led to an increasing need of investment, so that fiscal rules should be well designed in view of their potential side- effects.

The choice of the Golden rule as a case study is motivated by the consequences of the crises from the last decades that lead to strong imbalances between the European Union (EU) members. Consequently, fiscal discipline was jeopardized and the debate on the importance and usefulness of fiscal rules was also raised in the eurozone context where the SGP constitutes the supranational fiscal rule. As already discussed in previous chapters, the SGP's different reforms came with a growing complexity until the European Commission relaunched the review of EU economic governance on the $19^{th}$ October 2021 (initiated in February 2020 but suspended to focus on responding to the impacts of the pandemic crisis on the economic and social environment) to address the simplification of the fiscal framework[1]. The goal of this chapter is not to consider the peril fiscal rules and SGP as in Blanchard et al. [2020], not to propose an alternative solution for SGP re-

[1] See European Commission website and communication on EU governance review.

form as in Darvas et al. [2018], Debrun and Jonung [2019]. It aims at providing key highlights and warnings on computational issues for the net public investment which is the major component to adjust the Golden rule of public finance as potential candidate for the EU supranational fiscal rule. The famous Golden rule of public finance targets the public deficit excluding net public investment. Any change in the net public investment measurement would imply a change in the targeted value of the rule. Consequently, all statistical decisions to compute the net public investment may have consequences on the Golden rule target, its compliance as well as on policy recommendations (such as the European Commission recommendations during the European Semester). This work is inherent to the debate on fiscal rules performance as their performance assessment could be affected by the accuracy of national accounts indicators used as target of fiscal rules. Measurement errors in fiscal rules target could induce errors in judgment regarding the performance of fiscal rules and hence influence the debate on how to design them. Fiscal rules should therefore target indicators that are accurately computed.

The existing literature on the Golden rule mostly focused on the advantages or side-effects of its adoption for the economy (see, for an empirical assessment, Monperrus-Veroni and Saraceno [2005] that followed the approach of Eichengreen and Wyplosz [1998]; or Creel and Saraceno [2010] for discussion). This chapter does not follow this field and only considers the sensitivity of the measure of the Golden rule target to statistical/accounting assumptions. The United Kingdom (UK) experienced a Golden rule until 2009 and is the training case of this chapter. This analysis computes net public investment series to adjust to the public balance in the UK from 1998 to 2016. Despite the UK abandoned the Golden rule in 2009, the measure of the Golden rule target relies on statistical choices and not necessarily on the adoption of a Golden rule per se to handle the public finance. Indeed, the UK still follows a budget balance rule since 2010. In that sense, if public deficit is under control, changes in public deficit excluding net investment may be only affected by the statistical assumptions made in the chapter, and which influence the net public investment measure.

This analysis aims to highlight the importance of choices regarding the set of assumptions underlying the computation of CFC measures for the general government. CFC, also called Depreciation, is obtained in the process of capital stock measurement and this chapter compares all estimates to a so-called OECD benchmark series, which corresponds to the measures of CFC, net investment and budget balance that are obtained using the corresponding series for the general government sector sourced from the OECD national accounts database. The reader should note that on the OECD dataset, these series are available for the total general government sector, without any detailed on the underlying assets and industries involved. It is equally important to note that all series in the OECD national accounts database reflect the 2019 revision in the ONS National Accounts Blue Book, which revised the methodology underlying the calculation of capital stocks and hence CFC in the UK.

The CFC measurement challenges come with the capital stock measurement since, according to the OECD [2009], the capital stock is a crucial element in national accounts which allows to derive CFC. Capital stock is important since it allows to compute CFC. It also matters for productivity measurement as it constitutes an input for the growth accounting framework. And, last but not least, capital stock measurement relies on the GDP measure as CFC enters in its calculation. Any change in CFC of both public and private sector may affect the GDP (of the total economy). Net capital stock may be retrieved from company/business survey or through the direct application of the Perpetual Inventory Method (PIM)[2] which derives net capital stocks from the corresponding volumes of investment flows. As the CFC is a by-product of the PIM implementation, CFC is sensitive to methodological choices relating to the asset retirement pattern, asset prices, and the availability of long GFCF series which are essential to construct the initial capital stocks. It also depends on the age-price profile used to compute capital stock (OECD [2009]). By evaluating the changes in the depreciation pattern and parameters, we investigate how public CFC measurement may affect public net investment (which is defined as gross public net investment minus CFC) and thus the Golden rule targeted value. This chapter proposes a novel and original work by linking the statistical issues in general government CFC measurement and fiscal rules performance assessment. At this stage, the studies focusing on fiscal rules performance assessment did not question the accuracy of fiscal rules' targets. This chapter offers a novel insight in fiscal rules performance measurement by considering the importance of the method to calculate a particular target, namely the general government net investment.

The results may be decomposed as follow:

[2] PIM accumulates past Gross Fixed Capital Formation (GFCF thereafter), adjusted for retirement, and applies an age-price profile (for net wealth stock), or an age-efficiency profile (for productive capital stocks).

i) the use of different detailed levels of public assets and/or activities to compute general government CFC may imply variations up to 0.8% of GDP in general government net investment;

ii) we extend backwards general government GFCF series that are not long enough for this study and compute the initial capital stocks. At the beginning of the study period (that covers 1998-2016), some differences are observed between our series of derived CFC (from capital stock measures) and series retrieved from OECD; while they seem to be closer at the end of the study period. As the differences are not important and not persistent in the time, the backwards approach used in this chapter should be considered as a reasonable approach;

iii) the depreciation pattern (the age-price profile) is important for the CFC. Indeed, we compared a geometric approach (from author's computation) with:

a) two series using a combined normal retirement/straight-line profile. One comes from author's computation and the other one comes from the UK official source -namely the Office for National Statistics (ONS thereafter)- old methodology.

b) an age-price profile derived from a hyperbolic age-efficiency profile (sourcing from Office for National Statistics in the UK).

The changes in the form of the age-price function may involve changes in the general government net investment (which is the target of the Golden rule) up to 0.75% of GDP. Finally, the changes in the age-price profile parameters, in particular the depreciation rate, imply changes around 0.2% of GDP in the general government CFC measure.

The rest of the chapter is structured as follows. The section 2 exposes stylized facts on the Golden rule and net investment measures (including existing literature on capital stock measurement challenge) in the UK. The section 3 develops the empirical strategy, section 4 provides the results and section 5 proposes technical recommendations. Section 6 concludes the chapter.

## 4.2 Stylized facts on public net investment in the United Kingdom and the Golden rule

The chapter only focuses on the general government sector -general government assets and activities-[3]. The System of National Account (SNA; United Nations, 2008)[4] sets the assets classification and, following institutions as OECD, the analysis uses it as the reference for detailed assets level. The approach developed in this chapter, following the previous studies that paid attention to capital services measurement[5] (see e.g. Hall and Jorgenson [1967], Oulton [2007], Inklaar [2010]), assumes that we can perfectly define investment in each asset, of a certain firm, without error. In other words, it supposes that we perfectly identify the difference(s) between investment and expenses[6] and such distinction is performed upstream by institutions or statistical agencies that allocate "total investment in each asset across industries (rather than firms) and total investment by each industry across assets" (Inklaar [2010]).

Net investment is a national accounting core measure resulting from capital stock measurement series. Indeed, the age-price profile applied on Gross Investment allows to derive both capital stock and CFC. The net investment is strictly equal to Gross investment less CFC[7]. Such basic accounting definition is true for private and general government sectors. The general government net investment measure creates a link with the Golden rule that may ensure intergenerational equity by making "the cost of public expenditures be spread over time in a manner that reflects the intertemporal distribution of the benefits generated by those expenditures" (Robinson [1998]). Consequently, the Golden rule targets the general government net investment. There exists two famous versions of the Golden rule: the nominal or the structural[8]. The nominal Golden rule aims

[3] General government includes all government bodies, at all government levels (national, state, local) excluding public non-financial and financial corporations. Therefore, general government shouldn't be confused with public sector. For this reason, the chapter only refers to general government sector in the chapter. See European system of accounts (ESA, 2010) for further details on general government definition and activities.

[4] See Appendix 1 for further details on SNA assets classification.

[5] See Appendix 2 for an overview of capital aspects.

[6] See Inklaar [2010] for an illustration of such issue using research and development example which is often classified as an expense while it may be an investment.

[7] Translating the sentence in an equation gives: $Net\ investment = Gross\ Investment - CFC$.

[8] "Structural" refers to the cyclical adjustment of the budget balance to exclude the effects of the business

at ensuring the equilibrium in the general government nominal budget balance excluding net investment. The UK version of the Golden fiscal rule adopted in 1998 stipulated that the general government should borrow only to invest and not to finance current expenditure. Consequently, the general government balance excluding general government net investment, had to be balanced over a business cycle[9]. According to Creel and Saraceno [2010], the nominal budget balance corresponds to government "Net lending (+) or net borrowing (−)" and net investment is obtained using the government consumption of public capital as depreciation. In that sense, the nominal Golden rule (according to UK's definition) may be written:

$$\mathrm{GG\,BB} + (\mathrm{I} - \delta k) \geq 0 \tag{4.1}$$

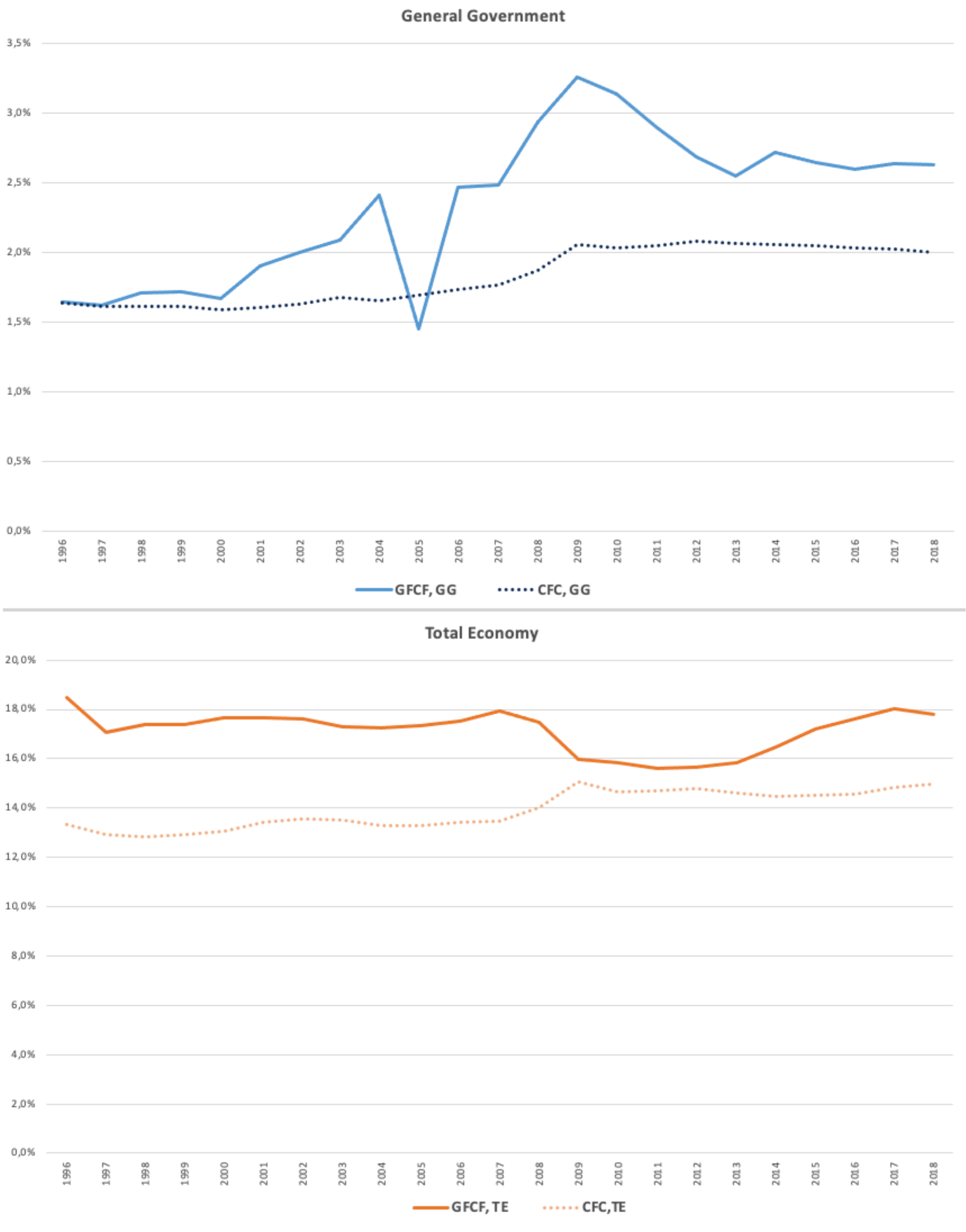
where GG BB corresponds to the general government Budget Balance, I is the gross general government investment (in percent of GDP), k is the general government capital stock (in percent of GDP) which depreciates at rate δ (which has no units). The alternative structural form of the Golden rule uses the structural public budget balance instead of the total one. Thus, the structural version may be written:

$$\mathrm{GG\,BB}_s + (\mathrm{I} - \delta k) > 0 \tag{4.2}$$

According to these definitions, it appears that the general government net investment measure is crucial for the Golden rule implementation. Since, in national accounting net investment is defined as the difference between GFCF and the CFC, GFCF is important for at least two reasons: i) it is the first term of the aforementioned accounting definition of the net investment; ii) the availability of historical GFCF series is important for the capital stock measurement from which CFC is derived.

cycle.

[9]In Creel and Saraceno [2010], the nominal case studied sets that the public balance excluding net investment should be superior to 0.03.




Note: GG stands for "General Government and TE stands for "Total economy". GCFC corresponds to "Gross Fixed Capital formation" and CFC corresponds to "Consumption of Fixed Capital".
Source: Author using OECD national account data.

Figure 4.1: **Gross Fixed Capital Formation and Consumption of Fixed Capital in the United Kingdom, by sector, in percent of GDP between 1996 and 2018**

Figure 4.1 presents the evolution of the CFC and the GFCF of general government in comparison with their evolution in the total economy, in the United Kingdom between

1996 and 2018[10]. It suggests that the general government GFCF is more volatile than the total GFCF, while the general government CFC remains stable over period as the depreciation of assets is not affected by cyclical events. As CFC seems not impacted by the economic cycles, making CFC responsible for changes in net investment series, implies introducing changes in the CFC calculations itself. This is, in particular, the statistical assumptions behind the methodology to compute the net capital stock and derive CFC that will affect CFC series and thus net investment.

Entering into the composition of general government GFCF by asset and sub-activities (also called "industries") types requires investigations and assumptions. Indeed, this detailed levels are important for stylized facts overviews and discussions, as well as for our practical implementation. Nervetheless, the GFCF series by detailed asset type are not readily available for the public sector in the OECD data sources[11]. The chapter follows the System of National Accounts (2008)'s assets classification which sets the asset boundary for fixed assets as they correspond goods and services which are used in production for more than one year[12].

To get only the GFCF at assets level for the general government sector our approach consists in computing the share of the total general government GFCF in total economy GFCF (equation 4.3) and apply these shares to GFCF (equation 4.4) in industries in which we can reasonably assume that general government sector is present[13]. This "share ap-

[10]For the purpose of evolution overview, the period of this graphic covers a broader than our study period.

[11]An inconsistency is also observed since the net capital stock for Dwelling in public sector is different from 0 while GFCF in dwellings for public sector is equal to 0.

[12]It breaks down the assets into different categories, namely: Dwellings, Other buildings and structures (which covers Buildings other than dwellings, other structures, land improvements), Machinery and equipment (which represents transport equipment, ICT equipment, other machinery and equipment), weapons systems, cultivated biological resources (which be decomposed in animal resources yielding repeat products, tree crop and plant resources yielding repeat products), Costs of ownership transfer on non-produced assets, Intellectual property products (Research and development, Mineral exploration and evaluation, Computer software and databases (Computer software, Databases), Entertainment, literary or artistic originals, Other intellectual property products.

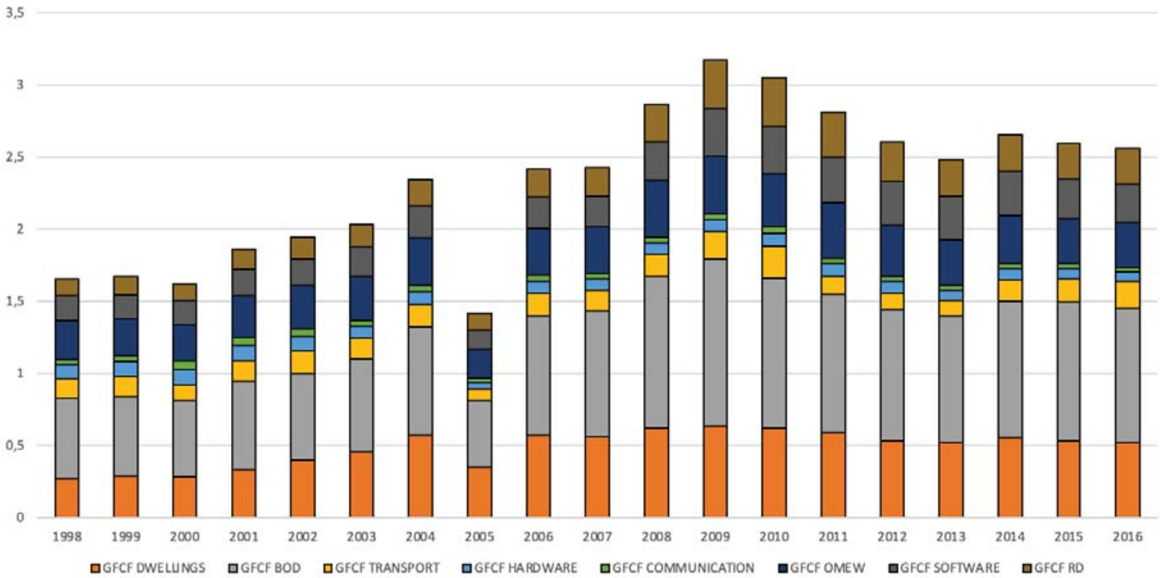
[13]This Chapter assumes that government operates in: Public administration; Administrative and support service activities; Arts, entertainment and recreation (ex: Museum); Education; Human health and social work activities; Other service and activities; Professional, scientific and technical activities; and in Real Estate to allow for GFCF in Dwellings. The selection of the industries in which, presumably, the public sector operates, is arbitrary. Further investigation about the actual correspondence between public sector activities and industries could help to improve this exercise in a future.

proach" is not applied when the industry/activity is Public administration which only covers public sector.

$$Share\ of\ \text{GG}\ sector = \frac{General\ Government\ \text{GFCF}}{Total\ \text{Economy}\ \text{GFCF}} \tag{4.3}$$

$$\text{GG GFCF}_{k,j} = Share\ of\ \text{GG}\ sector * \text{GFCF}_{k,j} \tag{4.4}$$

Where GG stands for "General Government", k represents each industry and I each asset (from SNA classification). This chapter assumes as strong assumption that all general government assets have the same share in total economy[14]. Nevertheless, this share is time-variant allowing adjustment in the time which may matter during crises where public sector may be increasingly active to provide fiscal support to economy.




Note: *See Appendix 1 for further details on SNA assets classification. GFCF stands for Gross Fixed Capital Formation. BOD = Building Other than Dwellings. OMEW= Other Machinery and Equipment and Weapons systems. RD= Research and development. Software and R&D are classified under Intellectual Property products in ONS. Hardware and Telecommunication equipment are classified as ICT equipment in ONS. CULT = Cultivated biological resources are equal to 0 in public sector. Variables are expressed in total GDP percentage. Author's series of general government GFCF is the result of applying general government share in total economy GFCF to each SNA asset total GFCF. Such procedure provides SNA assets general government GFCF measures, and their aggregation provide the total general government GFCF. See Section 4.3.1 for further details.
Source: Author based on OECD national accounts database.

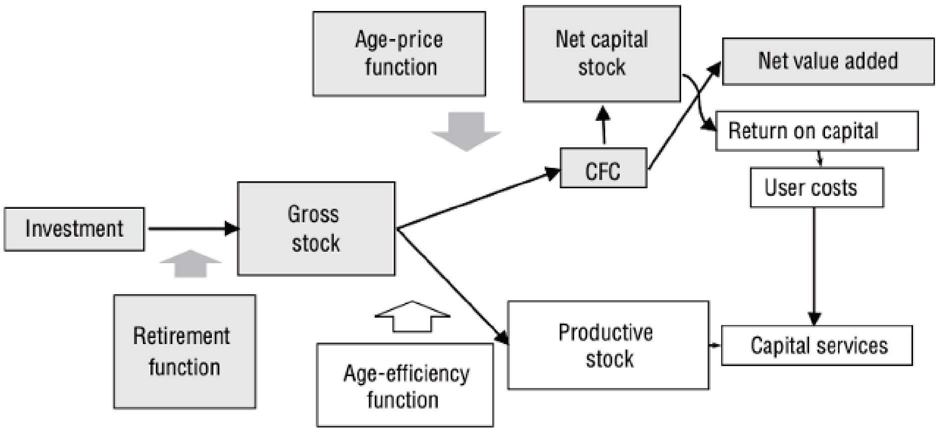
Figure 4.2: **Decomposition of General Government Gross Fixed Capital Formation in the United Kingdom, by Assets SNA boundary* type (in % of GDP) between 1998 and 2016**

[14]For example, if the share of general government GFCF represents 8% of total GFCF in a given year, the asset "R&D" will represent 8% of GFCF in the inductry "Professional, scientific and technical activities" for general government sector for this year.

Figure 4.2 shows that, all over the study period, Building Other than Dwellings (BOD) appears as the most important asset that may matter for capital stock. Such assets have long average service life including a maximum service life about 100 years for BOD according to the Office for National Statistics in the UK. Because the study covers less than two decades, assets with long service life will not disappear and assets as BOD still represent a large part of general government CFC at the end of the study period. On the other hand, Dwellings, that also have long service life, were multiplied by 3 at the end of the period. It seems that they benefit from a lot of investment and need a careful attention. On the other hand, Intangible assets as Research and Development represented a stable share of general government CFC, such as Software. This Chapter gives a particular attention to the assets that represent large part of total GFCF and that have long service life because they may intensively drive capital stock measurement (and thus CFC).

## 4.3 Empirical Strategy to compute general government net investment

The Measuring Capital OECD Manual (OECD [2009]) developed methodologies to compute CFC (depreciation level). The first one operates using the net capital stock and the second applied directly the rate of depreciation (OECD [2009]). Both of these, respectively, undirect and direct methods depend on the availability of investment series. The first approach is an undirect methodology because it applies the Perpetual Inventory Method (PIM) to retrieve the CFC. It is even more undirect when (instead of directly applying an age-price profile) it uses age-efficiency-profile to derive age-price profile for each asset types. The latter methodology guarantees the strict correspondence between age-efficiency profile and CFC. In practice countries use the Perpetual Inventory Method as they apply a depreciation function directly to gross series of assets investment asdescribed in 4.3.




Note : See also Appendix 3 for further details on Capital Stocks measurement based on SNA (2008).
Source: OECD (2009).

Figure 4.3: **Schema of Capital Stocks measurement**

As it is the most popular method according to OECD [2009], this chapter uses the PIM in the empirical strategy. To compute the CFC for a year, one needs the total change in the net capital stock between the end of year considered and the end of the previous year. For investment I (at constant prices) in industry k, asset i at time t, and using a geometric depreciation rate $\delta_i$, the capital stock A of each asset in each industry can be estimated using the PIM method as follows[15]:

$$A_{k,i,t} = (1-\delta_i)A_{k,i,t-1} + I_{k,i,t} \tag{4.5}$$

Nevertheless, this PIM may be applied following a set of major assumptions that may affect the results. The level at which the capital stock is computed may affect the results (APO/OECD [2021]). The Equation 4.5 above presents the most detailed as possible level to compute the capital stock since it is computed for each asset in the SNA asset boundary, in each industry. An alternative strategy is to use GFCF series cross classified by asset type only, without introducing the industry classification. Also, the age-price profile matters when applying the PIM. Indeed, in the case of the geometric profile, both retirement and depreciation profile are combined directly in the geometric function. When using hyperbolic or straight-line approaches, the retirement function is applied in upstream and there are a lot of candidates for retirement function as the Well-bull, the lognormal, the

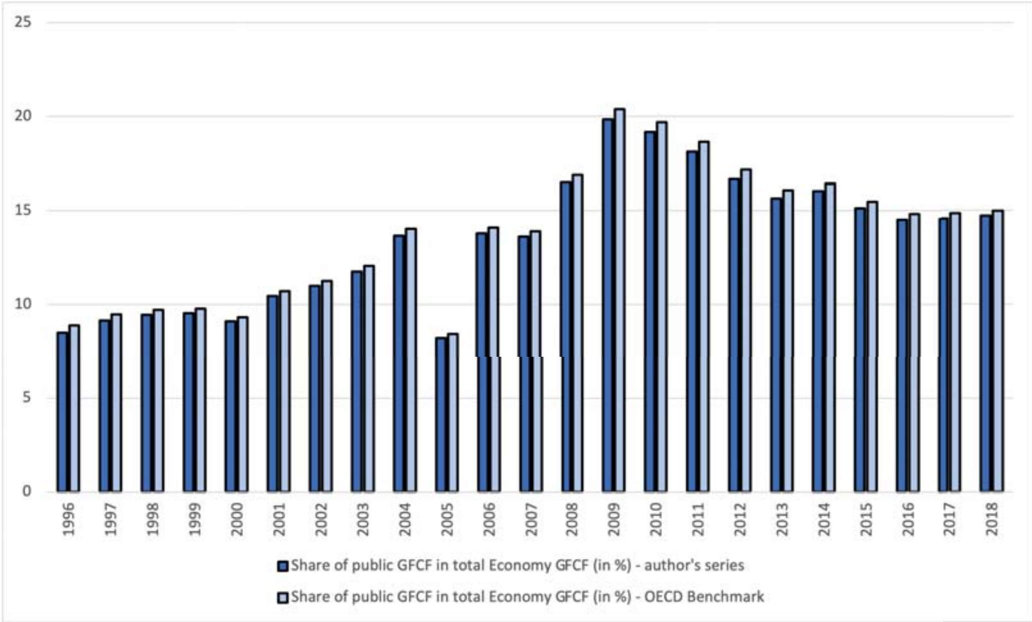
[15] Equation 4.5 relates PIM using a geometric model. See Section 4.3.2 for further details on age-price profile/depreciation pattern.

normal, the linear function (see OECD [2009] for an exhaustive list).

### 4.3.1 The sensitivity to the level of general government Consumption of Fixed Capital's calculation

According to the OECD [2009] and APO/OECD [2021] the accuracy of the total net capital stock and CFC series rely on the detail of the asset breakdown of investment series. The more disaggregated the asset breakdown, the more accurate the resulting net capital stock and CFC estimates. This study analyses the sensitivity of net capital stocks and CFC at different levels of aggregation across assets, namely i) it computes net capital stocks and CFC for each asset type in each different industry, which are later aggregated to obtain total net capital stocks and total CFC for the general government sector; ii) it computes net capital stocks and CFC for each asset type aggregated across industries and hence using an asset-specific depreciation rates common to all industries. As we have developed above, these tests are possible thanks to the hypothesis we have made with the equations 4.3 and 4.4.

We first must check if this approach does not introduce distortions in the GFCF series that we will use to calculate the CFC using the PIM. Figure 4.4 presents the share of the general government GFCF in total economy GFCF as sourced from the OECD national accounts database (4.3) and the share of general government GFCF in total economy GFCF calculated by the author as the aggregation of the general government GFCF shares in total GFCF in each given industry (each of them computed using 4.4).




Note: GFCF stands for Gross Fixed Capital Formation. Variables are expressed in total GDP percentage. Author's series of general government GFCF is the result of applying general government share to total economy GFCF for each asset in each general government industry. Such procedure provides general government GFCF by asset and activities measure and their aggregation provide the total general government GFCF.
Source: Autor and OECD national account database.

Figure 4.4: **Comparison of official series of General government Gross Fixed Capital Formation and series based on our public share assumption (in % of total GFCF), in the United Kingdom, between 1996 and 2018**

This figure 4.4 verifies the consistency of the approach described in equation 4.4 that uses share of the general government to retrieve the general government CFC. The method first applied general government share to total economy GFCF for each asset and each general government industry/activity and then aggregate them to get total general government GFCF. It is then compared with the OECD measure of general government CFC (sourced from official ONS series) of total public GFCF[16]. Figure 4.4 shows that the approach provides similar series over our study period 1998-2016[17]. In that sense, the methodology appears as a reasonable approach which looks comparable after aggregation at total general government level. A possible explanation for the lack of differences between our methodology and the official values may come from the low number of sub-sectors/sub-activities when looking at the general government sector. Thus, the assumptions made have fewer consequences than if one were to tackle a sector with many sub-industries, in

[16]OECD provides data on GFCF by sectors or by assets but not by sectors and assets. In that sense, total public GFCF is available but not public GFCF by asset type.

[17]For the purpose of evolution overview, the graphic covers a broader period than our study period.

particular the private sector. Our approach may therefore be reasonably valuable for the public sector but we cannot guarantee that it works for the private sector. Finally, we can apply this methodology to get the most detailed series of capital stock and CFC to test the sensitivity of the results to the computation level.

On the other hand, the figure 4.4 highlights that the general government GFCF ranges from 8% to 20% of the GFCF of the total economy. General government is not the major investor of the economy but may explain up to 1/5 of the variations of the GFCF of the economy. Consequently, we expect that changes in the measurement of the capital stock and the CFC, which are linked to past investment (past GFCF) flows through the PIM approach, can also significantly impact the most aggregate level (total economy level) of both capital stock and CFC.

### 4.3.2 The sensitivity of Consumption of Fixed Capital to the depreciation pattern and age-price profile parameters

#### Depreciation pattern

A key assumption in the measurement of capital stocks (and CFC) is the choice of the age-price (also named depreciation) profile. Such profile reflects the loss in value of a capital that aged[18]. Typically, the depreciation pattern that applies to a single asset of a cohort of assets is combined with a retirement function so to construct a so-called combined age-price/retirement profile[19].

OECD [2009] describes different approaches to estimate the combination of age-price function with retirement function: i) using empirical evidence on average service life of the assets and set up an assumption about the functional form of the depreciation pattern; ii) using information on used asset prices from second-hand asset markets and estimate depreciation using econometric approaches; iii) derive the age-price profile from the age-efficiency profile.

[18] Following APO/OECD [2021], it can be "illustrated by the pattern followed by the relative prices for different vintages of the same (homogenous) capital good".

[19] Indeed, all assets of a certain type (e.g. trucks) acquired in a given year constitute a cohort of assets. It is very unlikely that all the assets in a given cohort will retire or be discarded at the same age. For this reason, a retirement function is introduced, so to bring a distribution to account for different survival patterns.

In the first approach, the depreciation pattern may be assumed to follow different forms, namely straight-line or geometric. The straight-line (also called linear) model of depreciation imposes a constant amount loss in the asset value for every period. This constant loss is equal to 1/T where T corresponds to the service life of the asset. When using the geometric form, the asset value depreciates at a constant rate often noted as $\delta$. Obviously, the choice of the depreciation pattern and the parameters such as T or $\delta$ are key assumptions that will determine the accuracy of the series of capital stock and depreciation[20].

The second approach aims at supporting the assumptions made in the first approach. It uses price information on new and used assets to estimate depreciation. These studies often estimate a geometric depreciation rate (which is constant by definition) using second-hand asset prices. Fraumeni (1997) provided a survey on the topic and most studies concern the USA (Hulten and Wykoff [1981b], Koumanakos and Hwang [1988], Hall [1971], Grilliches [1960] or Jorgenson and Stiroh [1994]) but are also extended to other countries such as Canada (Baldwin et al. [2015]) or Japan (Suga [2018]). Both OECD [2009] and APO/OECD [2021] discussed the implications of these studies: i) different asset types have different age-price profiles[21]; ii) when plotting prices on the vertical axes and age on the horizontal one, age-price profiles typically present convexity towards the origin (Hulten and Wykoff [1981]). In that sense the straight-line approach seems not being the most accurate approach; but the geometric approach may be a reasonable approximation of the combined retirement profile/combined age-price profile (Hulten and Wykoff [1981], OECD [2009]).

The third approach consists in deriving the age-price profile from age-efficiency profiles.

From the 2019 revision of Blue Book[22], the Office for National Statistics in the UK uses

[20] See for official discussion in the Bureau of Economic Analysis in the US : https://apps.bea.gov/scb/account_articles/national/0797fr/maintext.htm

[21] “If price is plotted on the vertical axis and age horizontally, studies have found age-price profiles adopting a wide variety of functional forms (concave to the origin, horizontal lines, falling straight line and convex to the origin)” (APO/OECD [2021]).

[22] Every year the Office for National Statistics (ONS) updates the sources and methods for the UK National Accounts and publishes the latest estimates (including revisions to past periods) in the annual "UK National Accounts, the Blue Book" publication.

a hyperbolic age-efficiency profile to derive the depreciation pattern. The ONS previously used linear depreciation combined with a normal retirement function (Eurostat/OECD [2013] ). The normal distribution is famous in statistics applications. The frequency of the normal distribution is symmetrical and 95% of the probabilities are ranged around two standard deviations to the mean. Such property is not under debate regarding its usefulness for a retirement pattern. However, the Straight-line approach for the age-price profile supposes a constant depreciation. Such approach was not optimal since it does not properly reflect how certain assets depreciate. The straight-line depreciation method may do not accurately highlight the difference in the usage of an asset that may matter for some depreciable assets. In that sense, this study compares different series of the CFC and the net investment which were computed following different approaches. It starts from the from OECD as a Benchmark since it uses ONS official sources which corresponds to the most recent methodology of ONS. This analysis first uses a geometric profile as it appears as the most reasonable assumption following OECD [2009]. It then uses a straight-line to test how the change in ONS methodology, switching from a combined normal retirement/straight-line age price-profile to the new methodology described above, affected the net investment series.

**Depreciation rate and age service life measures**

As developed in the previous section, a convex age-price/retirement profile is obtained by combining an age-price profile for individual assets and a retirement profile for all assets in a cohort. Hulten and Wykoff [1981] argued in favor of using a geometric profile as reasonable approximation and there a geometric cohort depreciation rate.

In the absence of econometric estimates of geometric depreciation rates, δ has sometimes been estimated with the “declining balance method” and on the basis of information about average service lives. Hulten and Wykoff [1981] made the following suggestion for converting an average service life of a cohort into a depreciation rate. They propose a two-step procedure based on the “declining balance”:

$$\delta \equiv \frac{DBR}{ASL} \tag{4.6}$$

where DBR[23] is an estimated declining-balance rate and ASL the average age service

[23] Double declining balance rate imposes the strong hypothesis that efficiency of an asset is the same all

| Asset | Asset life used by country (Years) | | | | | |
|---|---|---|---|---|---|---|
| | United Kingdom | Germany | France | Netherlands | New Zealand | South Korea |
| Dwellings | 59 | 40 - 95 | - | 75 | 70 | 55 |
| Other buildings | 19 - 100 | 15 - 100 | 25 - 30 | 30 - 50 | 45 - 65 | 47 - 55 |
| Other structures | 19 - 100 | 25 - 150 | 60 | 25 - 55 | 25 - 110 | 30 - 65 |
| Land improvements | 19 - 100 | - | - | 1 | 30 - 58 | 17 |
| Machinery and equipment | 10 - 30 | 5 - 30 | 9 - 21 | 5 - 35 | 4 - 33 | 5 - 15 |
| Transportation equipment | 9 - 25 | 8 - 25 | 7 - 15 | 5 - 30 | 5 - 32 | 6 -30 |
| Computer software and databases | 5 | 5 - 30 [1] | 5 | 3 | 4 | 6 |
| R&D | 4 - 12 | 5 - 30 [1] | 10 | - | 10 | 9 -11 |

Source: Office for National Statistics, Blue Book 2018 UK, Table 1, UK assets service life compared to other countries.

Table 4.1: **Service life of SNA assets in 6 countries in the world**

life of the asset. A common approach us to use a double declining balance rate which corresponds to:

$$\delta \equiv \frac{2}{ASL} \tag{4.7}$$

Table 4.1, sourced from ONS, provides an overview of the assets service life used in different countries in the world. The UK assets service life reflect the assumptions made by the ONS in the UK before the revision in the Blue Book in 2019. This Chapter used these data since they were used until 2018 as the reference and it stops the analysis in 2016. As example, the “Other Buildings” have an average service life 3-4 times longer in UK than in France. However, it is clear that the service life differs considerably across countries. As a result, the use of the double declining balance assumption to estimate depreciation rates may introduce huge differences in depreciation rates for the same asset across countries and hence in capital stocks and CFC[24].

The changes highlighted in table 4.2 show that the ONS Blue Book 2019 introduced shorter average service life for assets. It implied a revision in capital stock estimates and

over its life but simplify the exercise in a presence of low data availability.

[24] See in particular OECD [2009] page 11-12 for further discussion.

| Asset description | Last estimation | Weighted old lives (years) | Weighted new lives (years) | New-old life (years) |
|---|---|---|---|---|
| Dwellings | Dean, 1964 | 59 | 50 | -9 |
| Other buildings | Dean, 1964 | 65 | 37 | -28 |
| Other structures | Dean, 1964 | 65 | 48 | -17 |
| Land improvements | Dean, 1964 | 65 | 20 | -45 |
| Transport equipment | Dean, 1964 | 11 | 15 | 4 |
| Telecommunication equipment | NIESR, 1993 | 9 | 18 | 9 |
| Computer hardware | Vaze, 2001 | 5 | 5 | 0 |
| Machinery and equipment | Dean, 1964 and NIESR, 1993 | 26 | 21 | -5 |
| Weapons systems | Based on other countries, 2014 | 20 | 20 | 0 |
| Cultivated Biological Resources | ONS | 10 | 6 | -4 |
| Computer software and databases | Vaze, 2001 | 5 | 5 | 0 |
| Entertainment, literary and artistic originals | Goodridge, 2008 | 15 | 10 | -5 |
| Research & Development | ONS, 2014 | 7 | 9 | 2 |
| Mineral exploration and evaluation | ONS | 10 | 15 | 5 |

Source: Office for National Statistics

Table 4.2: **Assets' life introduced in Blue Book 2019 vs old version**

thus in CFC. The ONS [2019] explained that the biggest impact they observed on net capital stock comes from the assets with long service life, such as “other buildings”, “dwellings” and “other machinery and equipment”. This supports the intuition introduced before: the fact that assets with long service life drastically affect CFC and net investment, and therefore need to be accurately retrieved. On the contrary “other structures” and “land improvements”, that were submitted to importany changes between the “old and new asset lives”, do not imply important changes in total capital stock estimates since they do not represent a big part of GFCF (ONS [2019]). For this reason, these asset types are excluded from the sensitivity analysis conducted in the present study.

Equations 4.6 and 4.7 make the link between depreciation measure and assets service life. Any change in the average service life implies a change in the depreciation rate when the DBR is kept constant. This study tests the impact of using estimates of average service life by asset type and industry as sourced from ONS. Average service lives for different assets across all industries correspond to the values presented for new service lives in Table 4.2. The study then tests how this change in average asset service life impacts estimates of net capital stocks, CFC, net investment and the budget balance of the general government sector.

### 4.3.3 The issue of the initial capital stock measurement

The series of GFCF used in this chapter starts in 1995 while some assets have long service life as dwellings. This is thus an issue to consider when computing capital stock using the Perpetual Inventory Method. To solve this issue, GFCF series are extrapolated using GDP growth rates before 1995. When long enough series of GFCF data are not available, historical GDP data should be used to extend backwards the GFCF series. Following growth theory, such approach assumes that the growth rates of GFCF and GDP are reasonably close. Historical GFCF data do not include exceptional increase or decrease in capital stock due to exceptional events such as wars or natural catastrophes that dramatically affect volumes of assets. The use of GDP growth to extend backwards GFCF series may help to capture the impact of events such on GFCF series. Such computational challenge is important for several eurozone countries since the length[25] of GFCF series is not the

[25]The longer of GFCF series should be compared to the average service life of the assets. If the first date of GFCF data availability is superior to the longer average service life across assets there is no particular issue

same across members. Countries such as Germany or France present long time series while some East European members don't. Consequently, the choice of the methodology to compute the initial capital stock may affect the accuracy of the net capital stock and CFC series. However, this aspect of capital stock measurement is not further discussed in the present study.

## 4.4 Results

This section exposes the results from the sensitivity tests. It presents the sensitivity of general government CFC, the general government net investment and the general government budget balance adjusted from net investment, to changes in:

i) The detail levels of the Gross Fixed Capital Formation series introduced in the Perpetual Iventory Method to compute the capital stock and retrieve the depreciation (CFC); by testing first the asset and activity breakdown and secondly the asset breakdown only;

ii) The depreciation pattern: by using alternatively a geometric profile for the entire cohort of assets and a combined normal retirement/straight-line profile[26] (one from our computation and the old official series from ONS before their revision in the Blue Book of 2019). These scenarios are compared to the ONS series that employed an age-efficiency profile to derive the age-price profile (reflected in OECD benchmark series);

iii) The depreciation rate: generated by changing the average service life of the assets. The series are computed using a geometric profile (as it is a reasonable approximation for the cohort of assets according to Hulten and Wykoff [1981], OECD [2009]) and the most disaggregated level (assets and activities level).

### 4.4.1 Sensitivity to the level of capital stock's calculation

In this subsection, all the series are computed using the geometric approach as recommended by OECD [2009]. The level to compute the capital stock and retrieve CFC changes. The first series follows OECD [2009] and APO/OECD [2021] recommendations by using assets and activities level. The second series is computed at assets level only. These two series are compared to the OECD SNA data series.

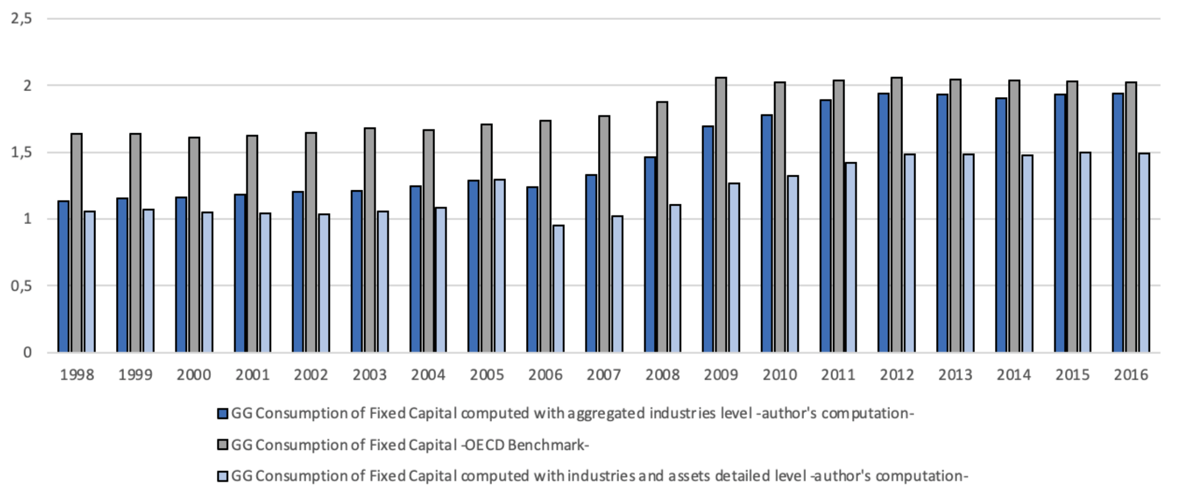
while it is important in the reverse case.

[26] The combination of a normal retirement with straight-line depreciation function corresponds to ONS old methodology.

Figure 4.5 highlights low differences between general government CFC series computed using the most detailed level (asset and activities detailed level) and the asset breakdown approach, at the beginning of the study period. The differences increase when assets aged at the end of period suggesting that the detail level at which CFC is derived matter when applying the depreciation profile. The most detailed series differ more from the OECD benchmark series. It confirms that the detail level of computation and the assumption on the general government industries to include in the analysis affect the results. As illustration of the importance of considering the detailed level of computation, differences up to 0.8% of GDP are observed in 2006, 2009 and 2010, between our most detailed series and the OECD Benchmark series. Nevertheless, at the end of the period the differences between the series are lower than at the beginning and they seem to follow a similar stable evolution over the period.

The series employing the asset breakdown differ from OECD data at the beginning of the period and get closer at the end of the period. This series appears less stable over time, suggesting that using less detailed level for CFC computation makes it less precise. The series showed that general government CFC represented 1.15% of GDP at the beginning of the period and 1.9% at the end of the period. It corresponds to 0.75% of variation of the series itself. We do not observe such variations in the two other series that do not increase by more than 0.5% over the period.

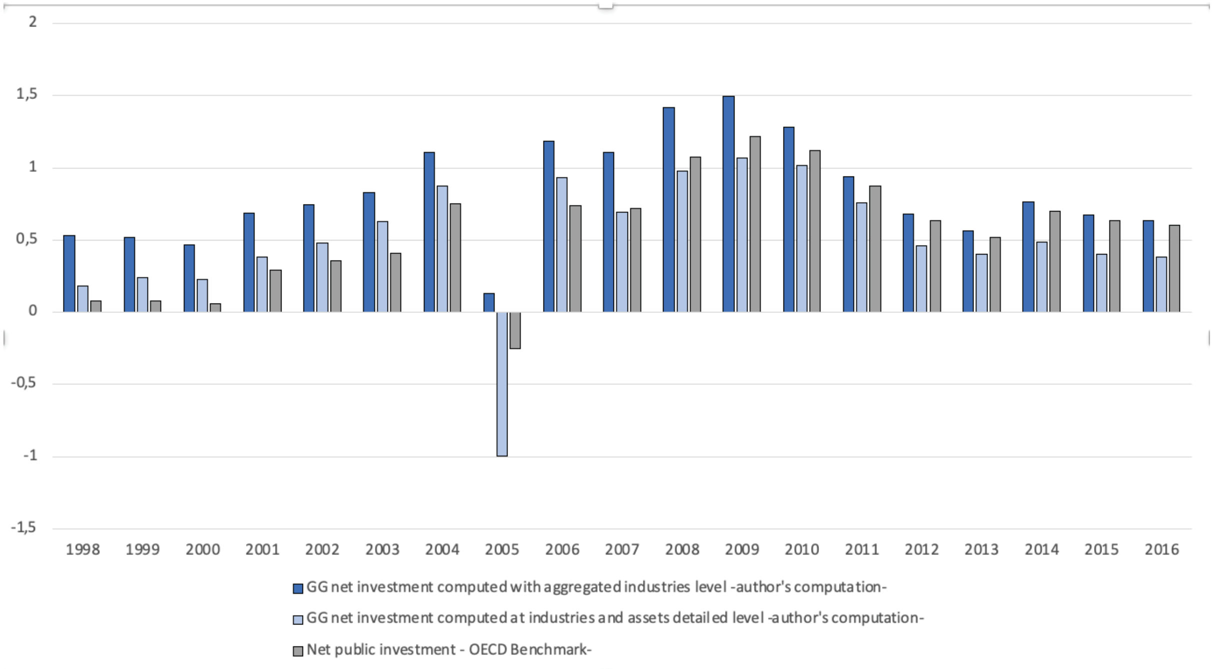
That's being said, the differences in the first years may be due to the initial stock computation. The step of choosing an approach for initial capital stock computation seems to be a major driver of the results. Indeed, we adopted a methodology of extending backwards the series based on the GDP growth rate as the GFCF series are available from 1995 while some assets have an average service life superior to 60 years in the old methodology which is problematic for PIM application. As the most detailed series does not much vary over time it couldn't perfectly fit the OECD Benchmark at the end of the period while the less detailed one, which varies more, reconciles the OECD Benchmarks series. The impact of the method to calculate the initial capital stock on the detailed series does not allow it to catch up with the official series, but it appears to be the series that best reflects the evolution of the CFC (figure 4.1 showed that the CFC did not vary much over time in general). However, it will take several years to reconcile a geometric approach with the hyperbolic approach (OECD Benchmark) when differences are observed in the initial capital stocks which are computed at the most detailed level. The use of a less detailed method, producing more volatile CFC, is not a precise choice although it seems attractive to reconcile series quickly.




Note: GG refers to General Government. OECD Benchmark series reflect ONS new methodology series.
Source: author.

Figure 4.5: **Comparison of General Government Consumption of Fixed Capital in the UK, using different computational levels (in % of GDP), in the United Kingdom, between 1998 and 2016**

Figure 4.6 confirms that the detail of the GFCF series breakdown (i.e. across assets and industries vs. across assets only) is important for the accuracy of general government net investment series. Consequently, the availability of detailed data and the availability of long time series of GFCF are essential to conduct accurate estimations for series that enter into the Golden rule target calculation.




Note: GG refers to General Government. OECD series is computed by deducting OECD series of general government CFC to general government GFCF.
Source: author.

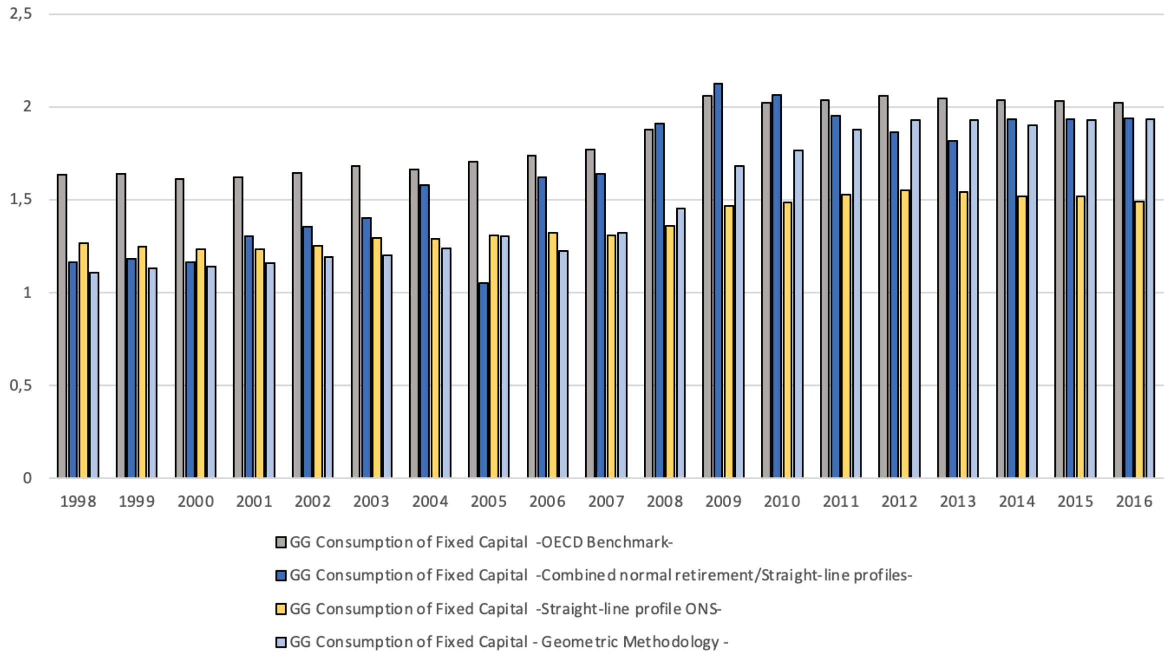
Figure 4.6: **Sensitivity of the general government net investment measure to the level at which net capital stock is computed (in % of GDP), in the United Kingdom, between 1998 and 2016**

Finally, we assume that some differences between our computed series and the official series (OECD Benchmark series) are due to the initial capital stock measurement as a result of GFCF backwards extension approach, in particular at the beginning of the study period. Although the most detailed series shows differences with the Benchmark series, it remains stable over time and reflects the same evolution as the official series. This is not the case for the less detailed level which fails to reflect the evolution of net investment in particular in 2005 and the evolution of the CFC is the most unstable while depreciation does not vary as much in the other series. As the less detail series may vary over time and produce more sensitive government net investment series, countries may consider the most detailed series of general government GFCF to compute the depreciation. This recommendation is in line with OECD [2009] and requieres that countries investigate a lot of efforts as some EU members do not aldready have such detailed data which might finally be problematic in the Golden rule assessment. Otherwise, when long time series of GFCF are not available for a country, the use of the backwards extension based on GDP approach may be a reasonnable approach as differences gradually disappear when moving in time.

### 4.4.2 Sensitivity to depreciation patterns/age-price profile

This section tests the sensitivity of the general government CFC, the general government net investment and the general government budget balance excluding net investment to changes in the depreciation pattern by using alternatively a geometric profile for the entire cohort of assets and a combined normal retirement/straight-line profiles. The latter scenario is represented in two series: one that we calculate ourselves and the former official ONS series that previously used this approach (reported as "ONS old methodology"). These methodologies are also compared to the approach that uses an age-efficiency profile to derive the age-price profile, according to the ONS current methodology and reported as "OECD Benchmark" series.

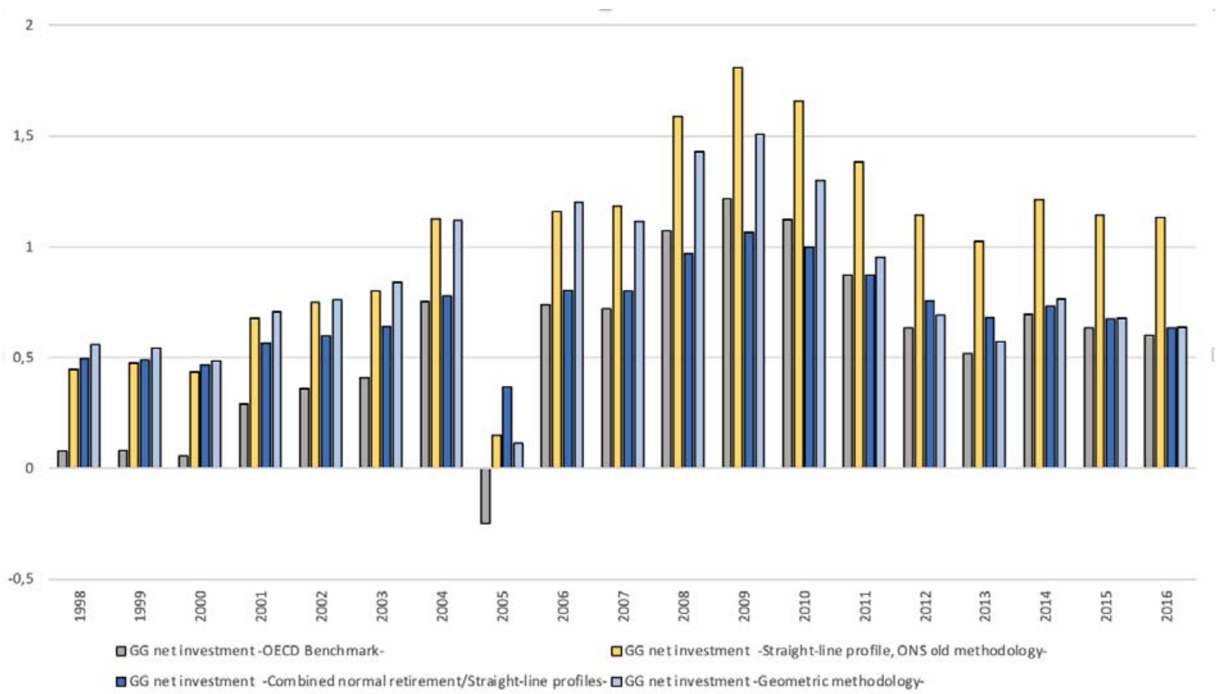
The Figure 4.7 suggests that the changes in the depreciation pattern imply high differences during economic stressful periods. Indeed, in pre-Global Financial Crisis and during the crises, the straight-line methodology does not adjust the series and looks higher than the CFC series obtained with a geometric pattern. Indeed, in pre-Global Financial Crisis and during the crises, the straight-line methodology does not adjust the series and looks higher than the CFC series obtained with a geometric pattern. The ONS old methodology seems to not appropriately capture the loss in value of assets as they age because it appears as approximatively constant in time while assets depreciate. Whereas the ONS new methodology (reflected by OECD Benchmark series) seems to be reflect better the lost in value of assets as it variates more over time. Our series that tried to reproduce the ONS old methodology does not fit with ONS old series. First, our series that used a combined normal retirement/straight line profiles increases until 2008 as it seems affected by our assumption on initial capital stock measurement. After 2008, the series are stable, reflecting the default of this straight-line approach which does not well reflect the loss in value of the assets. As the issue concerns the beginning of the period, it thus suggests that our approach to extend backwards GFCF series could not be used when a straight-line age-price profile is assumed.




Note: GG refers to General Government. ONS refers to Office for National Statistics in United Kingdom. The ONS old methodology employed a combined normal retirement/straight-line. According to OECD metadata in National Account table 14.A, data are sourced from “statistics reported to OECD by member countries in their answers to annual national accounts questionnaire” and thus retrieved the ONS new methodology series.‘ ‘OECD Benchmark” thus reflects the ONS current approach.
Source: author.

Figure 4.7: **Sensitivity of general government CFC to different depreciation pattern (in % of GDP), in the United Kingdom, between 1998 and 2016**

The Figure 4.8 follows the highlights from figure 4.7. The comparison of geometric and combined normal retirement/straight-line profiles lead to huge differences in general government net investment in the same periods as discussed above because of the changes in general government CFC. Nevertheless, one may observe that general government net investment from ONS, obtained by applying an age-efficiency profile to retrieve the depreciation (CFC) and thus net investment, leads to over-optimistic general government net investment series.

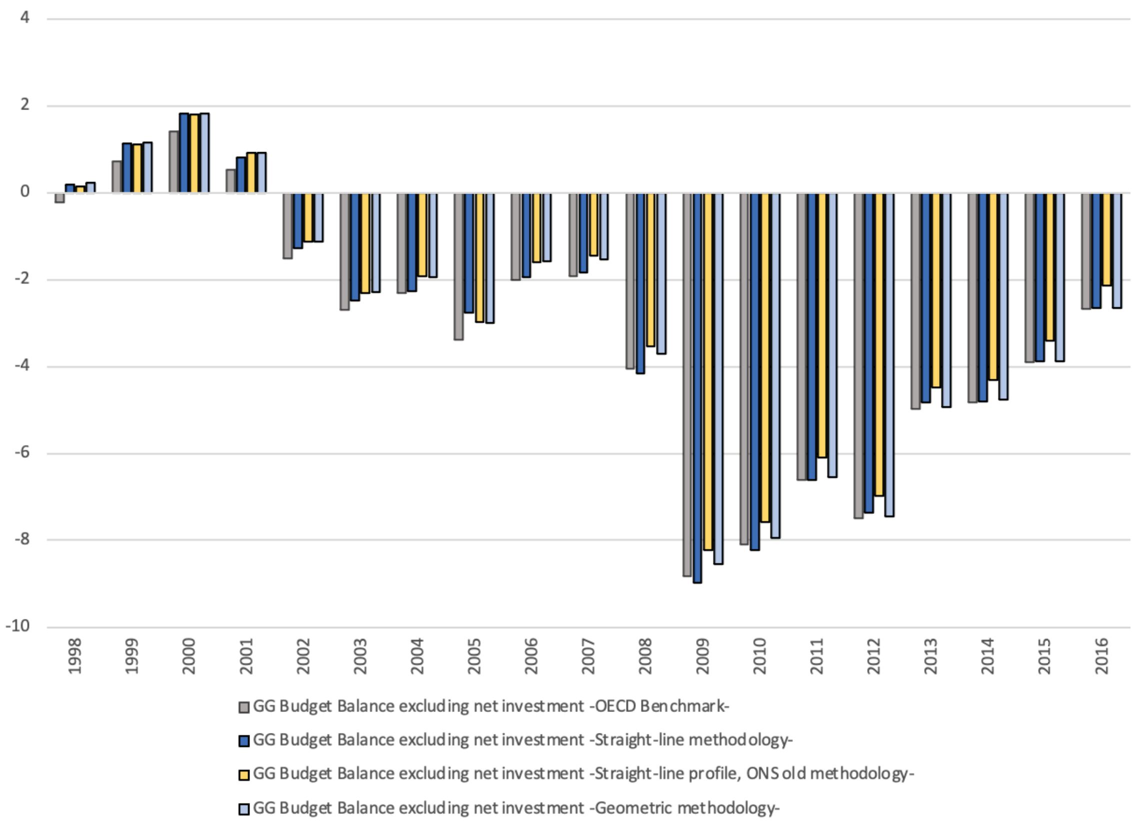



Note: GG refers to General Government. ONS refers to Office for National Statistics in United Kingdom. The ONS old methodology employed a combined normal retirement/straight-line. According to OECD metadata in National Account table 14.A, data are sourced from "statistics reported to OECD by member countries in their answers to annual national accounts questionnaire" and thus retrieved the ONS new methodology series. "OECD Benchmark" thus reflects the ONS current approach.
Source: author.

Figure 4.8: **Sensitivity of general government net investment to depreciation pattern (in % of GDP), in the United Kingdom, between 1998 and 2016**

Results from figure 4.8 highlight that the choice of the depreciation pattern is a strong determinant of the net investment. The differences in general government net investment are in general around 0.5% from one methodology to another (up to 0.75% between the ONS old methodology and and our series that tried to reproduce this old method from ONS between 2007 and 2012). First, differences look higher during economic turbulences as the most importance differences are observed during the subprimes crisis (2008-2009) and the sovereign debt crises (2010-2013). The general government GFCF was mostly volatile in these periods and methods that do not enough account for the depreciation seem to not properly reflect the depreciation reality. As a result, all fiscal plan and policy recommendations may be biased by such statistical lack in measurement accuracy which seem ever more important during economic crises. Consequently, the straight-line approach seems definitely to not be a reasonnable methodology. The differences between geometric approach and official series are still observed in the beginning of the period, suggesting that it is still the initial capital stock that plays a role in these differences. The geometric approach, which is simpler to implement than the ONS current methodology, may be a reasonable method to conduct.

Although it is difficult to observe the differences in the figure 4.9 (in particular because of the large variations in the general government balance excluding net investment on the vertical axe), the differences may be important as, for illustration, it concerns 0,61% of differences between the OECD Benchmark series and our computation using a straight-line profile, in 2005. Consequently, the depreciation pattern leads to differences in general government budget balance excluding net investment which is the target of the Golden rule of general government finance.




Note: GG refers to General Government. ONS refers to Office for National Statistics in United Kingdom. The ONS old methodology employed a combined normal retirement/straight-line. According to OECD metadata in National Account table 14.A, data are sourced from "statistics reported to OECD by member countries in their answers to annual national accounts questionnaire" and thus retrieved the ONS new methodology series.' 'OECD Benchmark" thus reflects the ONS current approach.
Source: author.

Figure 4.9: **Sensitivity of the general government Budget Balance adjusted from net investment to depreciation (CFC) measurement (in % of GDP), in the United Kingdom, between 1998 and 2016**

To have a more in-depth analysis of the differences observed in the Golden rule target according to the method used, table 4.3 provides the absolute value of the ratios be-

| | ratio "DDR/OECD Benchmark" | ratio "DDR/ONS old methodology" |
|---|---|---|
| Maximum | 1,76 | 1,60 |
| Minimum | 0,74 | 0,95 |

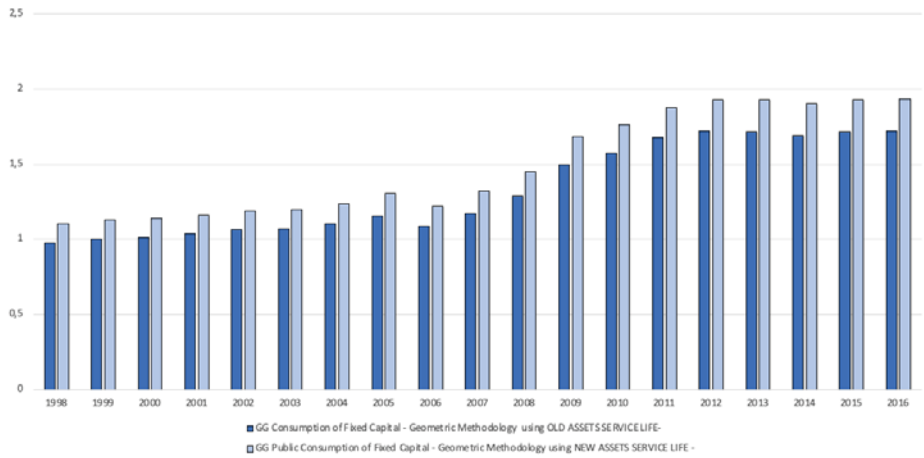
Note: OECD benchmark retrieves ONS official sources and uses the new methodology. It uses the most recent updates employing an age-profile derived from the age-efficiency profile. ONS series refer to ONS old series which employed a straight-line approach to compute general government capital stock and obtain general government CFC. DDR refers to the series that employed a double declining balance rate in a geometric age-price profile.
Source: author.

Table 4.3: **Differences in UK general government Budget Balance net of investment depending on age-price profile over 1998-2016 period**

tween general government balance excluding net investment using a geometric profile (with a Double Declining balance Rate (DDR) to compute the depreciation rate parameter), the OECD Benchmark series (which reflect the new ONS methodology) and the old ONS methodology. The differences seem important leading to a target of the Golden rule 1.76 times superior when using a geometric profile instead of the ONS new methodology. The changes in the ONS methodology implies a difference with the geometric approach that may be 1.6 times superior for the general government balance net of investment. The ratio of the series resulting from geometric approach over OECD benchmark seems to be more sensitive since the minimum of the ratio is 0.74 implying a higher dispersion of the differences between the series than when using the ONS old methodology which implies 0.95 as minimum of the ratio between the series. Finally, the changes in ONS methodology induced important consequences for net investment series and thus for Golden rule targeted indicator. These findings seem to support what the ONS observed, according to their Blue Book of 2019 ONS [2019], that is important changes in capital stock and CFC are due to assets with long average service life.

### 4.4.3 Sensitivity to changes in depreciation rate

This subsection conducts a sensitivity analysis to changes in the depreciation rate by changing the assets average service life. The ONS [2019] exposed the changes in the average service life introduced by the 2019 Blue Book of the ONS (ONS [2019]). All the series presented above use the depreciation rates at assets and activities level. The average service life of all assets across activities seems to match with the new assets service life exposed in Table 4.2. This testing exercise employs the old values of assets service life to induce change in the depreciation rate which still be computed using the double declining balance rate formula. The series presented are computed under a geometric profile.




Source: author.

Figure 4.10: **Sensitivity of consumption of fixed capital (in % of GDP) to changes in depreciation rate/assets service life**

For the space considerations we report only the results of the sensitivity analysis on general government CFC is reported on this section. To follow the recommendations of the OECD [2009] we used the most detailed series (assets and activities breakdown) because the differences observed between the series we computed in this section are only due to changes in the depreciation rate. Appendix 5 reports the results for the net general government investment. In figure 4.10, we observe that the new assets service life produce higher general government CFC. As highlighted in table 4.2, the new service lives are shorter implying that assets depreciate faster. Consequently, over the same periods, the new values of assets life imply higher depreciation/higher general government CFC. Moreover, Buildings Other than Dwellings (BOD) assets life have lost 28 years of lenght of life while they represent the higher share of the general government GFCF. Such a huge change in BOD may drive the capital stock measurement and thus general government CFC. This finding is in line with ONS [2019] which showed that the main changes in net capital stock after the modification of assets service life, are due to the assets with long service life as Dwellings or BOD.

However, the consequences are less significant than in our other tests since the variations do not exceed 0.2% in general. But, it is important to recall that the DDR method used to derive the depreciation rate is still simple and other more sophisticated formulations, such as econometrically derived depreciation rates, could produce different depreciation rates and induce a higher sensitivity of the CFC, and thus of the net investment.

## 4.5 Technical advices

Before entering into general government CFC measurement details, our first discussion concerns the extension of not long enough GFCF series. A country that faces short public GFCF series may reasonably decide to extent the GFCF series backwards making a plausible assumption about the long-run growth rate of investment. Following the Measuring Capital OECD Manual (OECD [2009]) a reasonable approach could be to use the of GDP growth rate whenever existing. According to growth theory, this approach supposes that GFCF grows at the same rate of the GDP, and that this rate is constant over time. This approach, may produce inaccurate estimates of capital stocks in the beginning of the period studied but errors gradually disappear when moving in time as they affect the estimate of the initial capital stock and fade away in each following year.

To produce concrete statistical recommendations, we first need to recall some key elements. In particular, the CFC (depreciation) -which is deducted from gross investment to get net investment- is necessarily derived from the net capital stock which depends on different factors; and in this study we were interested in the following: the level of asset breakdown in the implementation of the PIM to estimate the net capital stock; the functional form of the depreciation pattern; the depreciation rate. According to our results, table 4.4 proposes a qualitative assessment of the sensitivity of CFC to changes in the calculation of net capital stocks and, hence, CFC. This sensitivity analysis is used to guide our technical recommendations on the most appropriate methods to calculate the public CFC and derive the public net investment.

In this chapter we provided some evidence that the sensitivity of CFC is then strongly affecting net investment (i.e. gross investment minus CFC) and therefore the general government balance excluding net investment. Consequently, the performance of the Golden rule is also sensitive to these conclusions. A strong statistical framework in national accounts is thus of utmost importance for the fiscal performance, Furthermore, it also dictates the accuracy of the total economy GDP, a key national accounts indicator,

Table 4.4: **Qualitative assessment of the sensitivity of general government CFC to changes in capital stock measurement, and corresponding technical recommendations**

| Changes in capital stock computation | Sensitivity of general government CFC and recommendations |
|---|---|
| **Level of computation** | **Most sensitive (up to 0.8% change in percent of GDP)** |
| - Asset Breakdown (less detailed level tested) | Not recommended |
| - Assets and public industries/activities breakdown (most detailed level tested) | Recommended |
| **Depreciation pattern** | **Sensitive (up to 0.75% change in percent of GDP)** |
| - Straight-line profile | Not recommended |
| - Geometric profile | Most recommended (as a reasonnable assumption, in particular when no empirical supports are available) |
| - Hyperbolic profile* | Recommended with the implementation of statistical efforts to support this choice (it should well reproduce economic reality and be supported by econometric tests) |
| **Depreciation Rate** | **Less Sensitive than to other changes. Maximum variations induced in CFC are 0.2% of GDP**<br>*Warning: these conclusions are valuable ONLY for public sector, any generalization of recommendations for private sector cannot hold and all changes in average service life of the assets that affect their depreciation rate should be supported by econometric tests or business data survey. It should also be recalled that only the simplified Double Declining Balance Rate approach has been used and that other approaches (or even econometric results) could lead to a higher sensitivity.* |

Note: *Hyperbolic profile refers to the use of an age-efficiency profile (to obtain productive capital stock) to derive the age-price profile (to obtain net capital stock and then retrieve CFC).
Source: author.

as the CFC of public sector enters directly in the compilation of public sector's GDP and hence in total economy's GDP. In turn, changes in the CFC affects GDP, and as such, this has consequences on fiscal rules targets expressed in percentage of GDP and affects the performance of other fiscal rules than the Golden rule. More broadly, it impacts the assessment of the economic development, cross-country comparisons, and the measurement of potential-output. The more sensitive the CFC to changes in a given technique, the more rigorously the choice of that technique must be justified. Methods to compute

CFC that do not reflect economic reality (as the straight-line age-price profile) should be discarded, and the average assets service life should be estimated using strong empirical approaches. Indeed, too long average service lives for assets could carry over errors for decades.

## 4.6 Conclusion

This chapter proposed a sensitivity analysis of the targeted value of a Golden rule, namely the general government budget balance excluding general government net investment, which receives an increasing attention due to the need for general government investment to face the pandemic crisis. This Chapter does not discuss the relevance of the Golden rule for the EU context but it discusses the statistical issues that can considerably shape decision making on general government spending when this is framed on a Golden rule, and addresses corresponding recommendations for general government net investment measurement. This chapter is inherent to the debate on fiscal rules performance in the sense that fiscal rules' performance may be also affected by the accuracy of national account series used as target of fiscal rules.

We first tested how the detail of the breakdown of GFCF series at which capital stock is computed matters for the capital stock level and CFC in general government sector. In line with OECD [2009] and APO/OECD [2021], this study compared the impact of using an asset/industry breakdown and an asset breakdown only of GFCF series and found important differences in the series. Nevertheless, we observed differences between the chapter's series and OECD series (which uses the official UK sources, namely the Office for National Statistics) that may be first driven by the methodology to compute initial capital stock. The issue that concerns initial capital stock measurement comes from insufficiently long GFCF series. Consequently, countries that want to produce precise net capital stock, CFC and thus most precise series of general government net investment may need long series of GFCF at assets level following the SNA asset classification.

However, the CFC measure is highly sensitive to the depreciation rate and the form of the combined retirement/age-price profile. As the United Kingdom switched from a Straight-line depreciation form to a depreciation form derived from a hyperbolic age efficiency profile, it was interesting to see the differences that such change may imply.

We compared these approaches to the geometric approach recommended by the OECD [2009]. In this chapter we used a simple double declining balance rate (2/T with T the service life of the asset) to compute the depreciation rate used in the geometric approach. We are aware of the limits of such approach that imposes no change in efficiency or productive capacity as the asset changes. Nevertheless, this chapter was not an attempt to produce the most accurate series of net investment for the UK but tried to highlight how any change in the methodology may affect the series of general government CFC and net investment.

The depreciation pattern matters for the computation of the general government CFC and the net investment. We showed that the use of the depreciation rate in a geometric profile for a cohort of assets lead to different results than the OECD and the ONS old series. These results may be (at maximum) 1.76 times superior to OECD series (which means ONS new methodology), and 1.60 times superior comparing to the old methodology. Moreover, the comparison of a geometric profile to a combined normal retirement/straight line profile confirmed that the choice of the depreciation pattern affects the general government CFC series. Indeed, the use of normal retirement/straight-line age-price profiles produced huge differences in the general government CFC series during a period of bad economic conjuncture. This finding appears as important since general government net investment is affected in a similar way. As general government budget balance excluding net investment corresponds to the target of the fiscal Golden rule, any mistake in its assessment during economic crises may misinform recommendations for economic recovery. In that sense, all choices made in the Perpetual Inventory Method, including the depreciation pattern should be supported by empirical evidence and assessment.

On the other hand, general government CFC is also sensitive to changes in the depreciation rate induced by modifications in assets service life. Shorter service lives increase depreciation rates and thus increases the CFC. This result is even more important when the changes affect assets with long service life and which represent a large share of general government GFCF such as Building Other than Dwellings. Any countries should conduct empirical evidence of the assumptions engaged and justify the relevance of any change in the methodology to ensure the accuracy of net investment series that enter to the target of the Golden rule of public finance. Economists and statisticians from institutions in charge of fiscal monitoring may also conduct such empirical analysis to provide statistical recommendations.

The implementation of a Golden rule requires efforts in the countries national accounting statistic methodologies to conduct before its application. Implementing such a rule before better understanding the differences in methodologies across countries to estimate capital stocks can undermine any assessment of the Golden rule performance and lead to errors in judgements. Indeed, a country could be seen as badly (highly) disciplined regarding the Golden rule target whereas the general government balance excluding net investment is affected by the methodology underlying the computation of the depreciation. Finally, it is also important to remember that these results are important for the calculation of the GDP (since the CFC is included in its calculation), and therefore for many indicators derived from GDP and/or are linked to it.

# Appendices

**Appendix 1: System of National Account classification of Assets**

| 2008 SNA code | Produced fixed assets |
|---|---|
| **N111** | **Dwellings** |
| **N112** | **Other buildings and structures** |
| **N1121** | **Buildings other than dwellings** |
| **N1122** | **Other structures** |
| **N1123** | **Land improvements** |
| **N11M** | **Machinery and equipment and weapons systems** |
| **N1131** | Transport equipment |
| **N1132** | ICT equipment |
| **N11321** | Computer hardware |
| **N11322** | Telecommunications equipment |
| **N11O** | Other machinery and equipment and weapons systems |
| **N115** | Cultivated biological resources |
| **N117** | Intellectual property products |
| **N1171** | Research and development |
| **N1172** | Mineral exploration and evaluation |
| **N1173** | Computer software and databases |
| **N1174** | Entertainment, artistic and literary originals |
| **N1179** | Other intellectual property products |

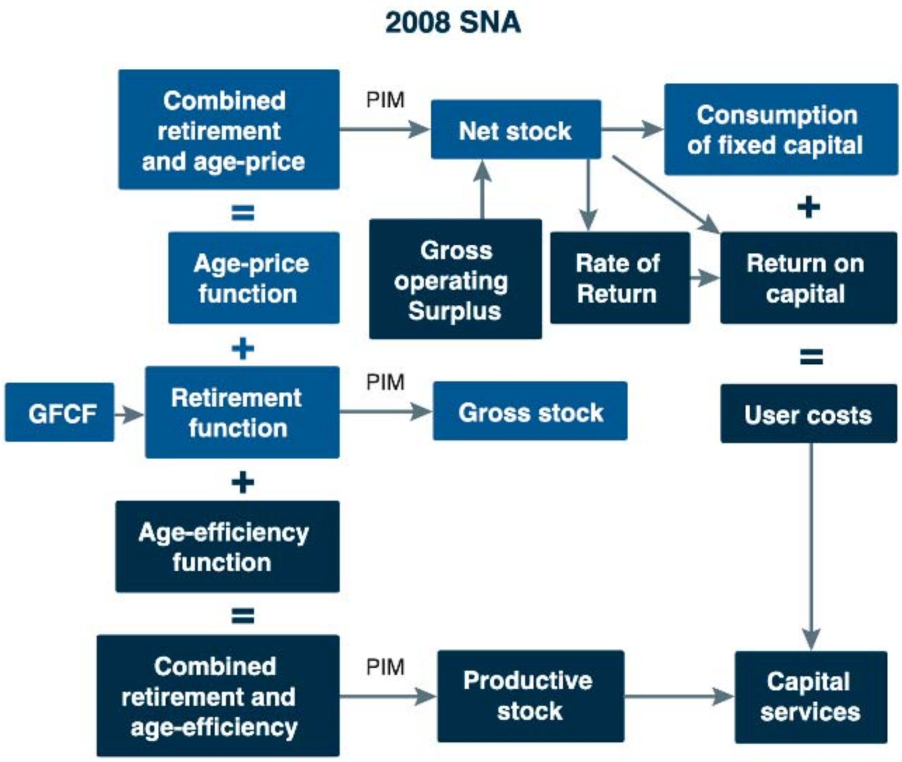
Source: APO/OECD (2021)

**Appendix 2: Capital Stocks aspects**

| | Income and wealth perspective | Production and productivity perspective |
|---|---|---|
| **Basic flow** | Investment | Investment |
| **Aggregation across assets of different age based on** | Depreciation profile (Age-price profile) | Age-efficiency profile |
| **Resulting stock for each class of assets** | Net capital stock by asset type | Productive stock by asset type |
| **Derived flow** | Depreciation (Consumption of Fixed Capital) | Capital services by type of asset |
| **Aggregation across different classes of assets based on** | Market prices | Price of capital services |
| **Resulting stocks** | Total net capital stock | Productive stock for each type of asset |
| **Derived measures** | Balance sheet entry, national wealth, net measures of income | Capital services, multifactor productivity |

*Source: OECD-APO (2021) Adapted from the Measuring Capital Manual: Second Edition (OECD, 2009)*

**Appendix 3: Schema of Capital Stocks measurement in System of Nantional Account (SNA, 2008)**




Source: Office for National Statistics

## Appendix 4: The comprehension of public sector activities

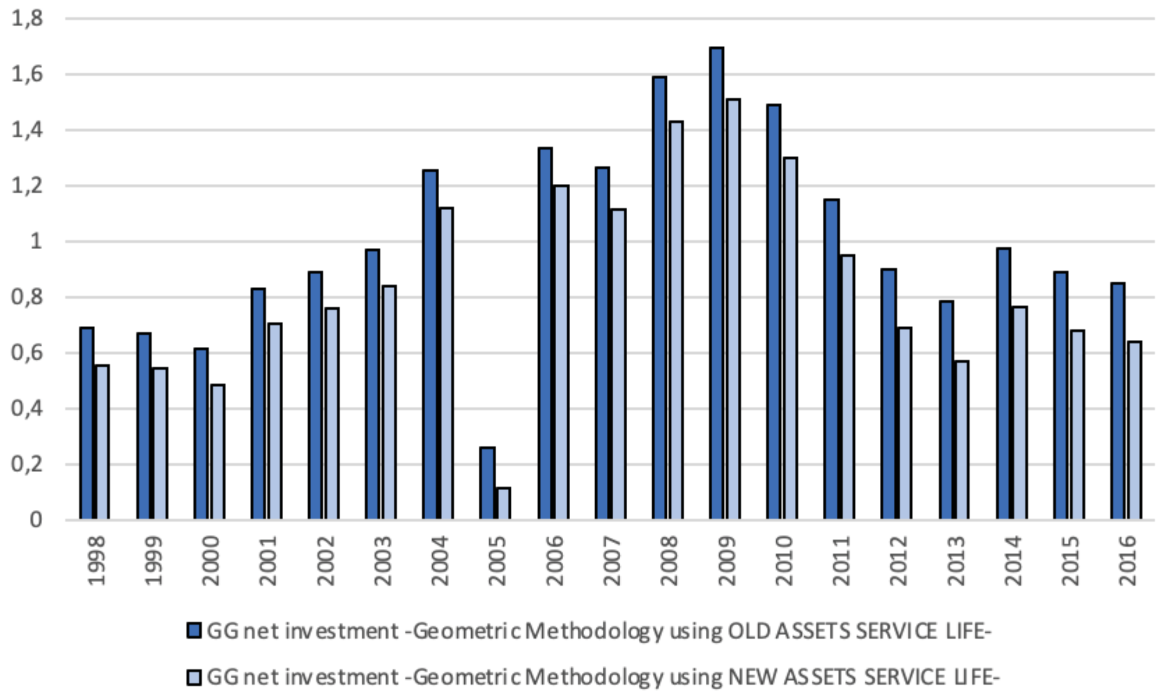
The OECD (2009) proposed the following classification of activities:

| ISIC Tabulation Categories | Description |
|---|---|
| A + B | Agriculture, hunting, forestry and fishing |
| C | Mining and quarrying |
| D | Manufacturing (*with 4 or 5 important activities separately identified*) |
| E | Electricity, gas and water supply |
| F | Construction |
| G + H | Wholesale and retail trade, repair of vehicles and household goods, hotels and Restaurants |
| I | Transport, storage and communications |
| J + K | Financial intermediation, real estate, renting and business activities |
| L | Public administration, defence and social security |
| M, N + O | Education, health and social work, other community, social and personal service activities |

Source: OECD (2009)

At the same, in page 151 the OECD Capital Manual (2009) explained that "In practice, a fully developed dataset may not be available. Also, the distinction between market and non-market producers is sometimes difficult to draw, specifically in industries such as education and health services where both types of producers operate. A simplified approach consists in combining all industries that are dominated by market producers into the "market sector", possibly with the exception of the real estate activities where provision of owner-occupied housing should be separately identified as production by households. The government sector would then be identified with public administration and defence (ISIC category L) and other community, social and personal services (ISIC category O)."

## Appendix 5: Sensitivity of general government net investment to changes in depreciation rate




Source: Authors using geometric pattern.

# REFERENCES

## Bibliography


APO/OECD. The sensitivity of capital services measurement: measure all assets and the cost of capital. *OECD Publishing*, 2021. 180, 181, 183, 184, 189, 201

J. Baldwin, L. Huku, and M. Tanguay. The Measurement of Quality Changes from Vintage Price Data." In Price Indexes and Quality Change. . *The Canadian Productivity Review*, 2015. 184

O. Blanchard, A. Leandro, and J. Zettelmeyer. Redesigning the EU Fiscal Rules: From Rules to Standards. *Mimeo, contribution presented to the 72nd Economic Policy Panel Meeting*, 2020. 170

J. Creel and F. Saraceno. European fiscal rules after the crisis. *Journal of Innovation Economics and Management*, pages 95–122, 2010. 171, 175

Z. Darvas, P. Martin, and X. Ragot. European Fiscal Rules Require a Major Overhaul. *Notes du conseil d'analyse economique*, 47, 2018. 171

X. Debrun and L. Jonung. Under threat: Rules-based fiscal policy and how to preserve it. *European Journal of Political Economy*, 57, 2019. 171

B. Eichengreen and C. Wyplosz. The stability pact: More than a minor nuisance? *Economic Policy*, 1998. 171

Eurostat/OECD. Eurostat-OECD, survey of national practices in estimating net stocks of structures. 2013. 185

Zvi. Grilliches. The demand for a durable input: U.s. farm tractors, 1921–57. in the demand for durable goods, edited by arnold harberger. *University of Chicago Press*, 1960. 184

R.E. Hall. The Measurement of Quality Changes from Vintage Price Data." In Price Indexes and Quality Change. . *Harvard University Press*, 1971. 184

R.E. Hall and D.W. Jorgenson. Tax Policy and Investment Behavior. *American Economic Review*, 1967. 174

C.R. Hulten and F.C. Wykoff. The estimation of economic depreciation using vintage asset prices: An application of the Box-Cox power transformation. *Journal of Econometrics*, 1981. 184, 185, 189

C.R. Hulten and F.C. Wykoff. he Measurement of Economic Depreciation. In Depreciation, Inflation, and the Taxation of Income from Capital, edited by Charles R. Hulten,. *The Urban Institute Press*, 1981b. 184

R. Inklaar. Towards Improved and Comparable Productivity Statistics: A Set of Recommendations for Statistical Policy. *The review of Income and Wealth*, 2010. 174

D. W. Jorgenson and K. Stiroh. Computers And Growth. *Economics of Innovation and New Technology*, 1994. 184

P. Koumanakos and J.C Hwang. The forms and rates of economic depreciation: The canadian experience. *Paper presented at the 50th anniversary meeting of the Conference on Research in Income and Wealth*, 1988. 184

C. Mathieu and H. Sterdyniak. Do we need fiscal rules? *Revue de l'OFCE 2013/1 No 127*, pages 189–233, 2013. 169

P. Monperrus-Veroni and F. Saraceno. Reform of the SGP: reducing or increasing the nuisance? *OFCE Working Paper*, 2005. 171

OECD. Measuring Capital, OECD Manual 2009: Second edition. *OECD Publishing*, 2009. 172, 179, 180, 181, 183, 184, 185, 186, 189, 192, 198, 199, 201, 202

ONS. National Accounts articles: Changes to the capital stock estimation methods for Blue Book 2019. 2019. 188, 197, 198

N. Oulton. Ex Post Versus Ex Ante Measures of the User Cost of Capital. *Review of Income and Wealth*, 2007. 174

M. Robinson. Measuring Compliance with the Golden Rule. *Fiscal Studies*, 19:447–462, 1998. 174

Y. Suga. Measurement of depreciation rates using microdata from disposal survey of japan. 2018. 184

# General conclusion

This thesis tackled a major economic and political challenge, namely the study of fiscal rules performance. The last decades have been scarred by deep economic crises which have led to significant indebtedness, resulting from accumulated public deficits. In each crisis, the fiscal rules, responsible for ensuring fiscal discipline alongside other complementary tools of states' fiscal framework, were debated. This thesis provided empirical evidence on the advantages and limits of fiscal rules performance which relates to different aspects of fiscal rules application, including: presence, compliance and enforcement (side)-effects, as well as the importance of the accuracy of their targeted indicators. This empirical thesis offers different methodologies to address the problem of fiscal rules performance assessment and alternatively uses standard econometrics, Machine Learning and statistical applications. In this way, this thesis proposes the introduction of Machine Learning methods to fiscal policies issues. It used both forecasting Machine Learning and causal Machine Learning showing that there is an existing and growing methodological field that may perform in fiscal performance monitoring and assessment.

Chapter 1 considered the fiscal rules performance regarding the effect of national fiscal rules adoption on fiscal discipline proxied by the popular CAPB as well as a novel measure of fiscal discipline (namely, the Global Financial Performance Index – GFPI). This chapter employed a propensity score matching method to account for potential endogeneity and showed that the fiscal rules significantly improves the GFPI, corroborating their favorable effect on the popular CAPB measure of fiscal discipline. This effect may be affected by the type of fiscal rule and different structural factors such as countries' characteristics and rules' design.

Chapter 2 discussed the issue of compliance with fiscal rules. In general, a poor compliance track record is observed. This is particularly important in the context of the Euro- zone, a common currency area, where the soundness of public finances may be even more vital (Mundell [1963]). This chapter proposed a tool to consider in the monitoring of fiscal rules performance as it helps to forecast it. It could be interpreted as a "risk-management" approach for fiscal surveillance applying a Machine Learning forecasting approach to the issue of fiscal rules compliance.

We highlighted that Support Vector Machine models consistently outperform the standard logistic regression and noted that the MIP scoreboard indicators (first and secondary indicators) are not sufficient to forecast the 3% limit of SGP compliance and hardly prevent imbalances. We identified the key features to forecast the SGP's compliance, and created a SVM model providing an accuracy ranking between 90.4% and 98.1% of performance of the prediction for compliance with SGP.

Chapter 3 investigated the potential undesirable effects of fiscal rules enforcement by assessing national budget balance rules compliance side-effects on social welfare. It tackled the limits of fiscal rules performance to consider their side-effects when designing them. It used Double/Debiased Machine Learning methodology, alternatively using LASSO and Boosting feature selection algorithms as robustness tests and the results are consistent across these methods. From the feature selection (first step of the approach), a set of ten key determinants for BBR's compliance are identified including voter preferences suggesting that voter preferences need to be taken into account in fiscal rules analyses as recommended by Wyplosz [2012]. Finally, average treatment effect results (second step) showed that achieving compliance with national Budget Balance Rules seems to generate side-effects because governments reallocate their spending: Governments favor public Gross Fixed Capital Formation over Government Final Consumption, which includes social spending. Consequently, BBR's compliance may be detrimental for social expenditure and have an increasing effect on inequalities, with the greatest effects on the poorest groups. These empirical findings supported evidence of side-effects from fiscal rules strict compliance, but when the definition of the fiscal rules' compliance is relaxed, the compliance may support economic growth. In that sense, introducing flexibility in the fiscal rules compliance definition and considering this flexibility when designing fiscal rules may limit the side-effects of fiscal rules' compliance. Consequently, we do not recommend an abandonment of fiscal rules and their rigorous application, but we recommend an increased focus on their design, in particular by considering flexible rules and ensuring a balance of both social spending and productive spending.

Finally, Chapter 4 entered into the debate on fiscal rules performance which may be affected by fiscal rules' design, measurement and optimality. We know that it is not possible to implement an "ideal" (or optimal) fiscal rule according to Kopits and Symansky (1998)'s definition, since there exists an insurmountable trilemma in the design of fiscal rules (Debrun and Jonung [2018]). In that sense, we explored the literatures propositions regarding the implementation of better fiscal rules which could be considered second-best options. Chapter 4 focused on the Golden rule (that targets the public budget balance excluding public net investment) which received increasing attention due to the need of public investment to face the pandemic crisis. Our work analysed the statistical issues and addressed corresponding recommendations for public net investment measurement since the quality of the targeted indicators of fiscal rules matters for their performance assessment. The chapter concluded that the implementation of the Golden rule of public finance needs strong statistical efforts and transparency before it enters into force. Putting such a rule in place without first justifying and reinforcing the methodologies of countries' national accounts, in particular to compute consumption of fixed capital (depreciation), could be perilous for the assessment of the Golden rule performance and lead to errors in judgments. We do not conclude that it is impossible to find a reasonable second-best optimal fiscal rule that can achieve fiscal discipline without damaging public productive investment. A Golden rule may relax the pressure on social spending since a part of public spending (the productive spending) is no longer taken into account. But, the implementation of such second-best rules is possible only with rigorous statistical measurement, institutional alignment and countries accounting transparency. These considerations of public consumption on fixed capital also impacts GDP measurement (which includes the consumption of fixed capital), with knock-on effects for several economic indicators (e.g. potential output, output gap, and any variables expressed as a percentage of GDP).

This research may be extended in several ways. First, and according to the last chapter, it may be possible to assess the accuracy of the measure of public net investment in countries interested in adopting the Golden rule such as in the Euro zone. Chapter 4 used the UK as case study, but it may be extended to other countries, in particular countries that did not have long series of public GFCF. Further sensitivity analysis could be conducted, such as alternative depreciation rates since Chapter 4 assumed the double declining balance rate. Other estimations and formula to compute initial capital stocks may be used to retrieve consumption of fixed capital. Any changes in general government consumption of fixed capital measurement and value may affect the composition of GDP which is not taken into account in our analysis which assumes a constant GDP to only observed changes net investment itself than in its measurement as percent of GDP.

The other chapters also leave room for future research. Chapter 1 may be extended by looking at the response of fiscal discipline to the so-called second-generation fiscal rules (see Caselli et al. [2018]), which add flexibility and enforceability to the simplicity feature of traditional fiscal rules. Chapter 1 was constrained to EU countries and future studies could investigate the effect of fiscal rules on fiscal discipline (defined by a larger indicator than CAPB as Global fiscal performance index) in other common currency areas such as the two African monetary unions – the CEMAC and the WAEMU. The monitoring tool provided by Chapter 2 to forecast fiscal rules compliance could be extended to other rules included in the SGP and which are summarized in the Compliance Tracker Database (Larch and Santacroce [2020]), such as the structural balance rule and expenditure rule. The model could also be transposed to forecast national fiscal policy outcomes using large datasets at national level. Chapter 3 may also be extended to the application of Machine Learning in other areas of fiscal policy area. Other national fiscal rules effect may be explored using the Double/Debiased Machine Learning approach such as supranational and/or subnational rules. On the other hand, the budget balance rules compliance database constructed for the study may be powered by other budget balance rules on a longer period even if they are still not applied to allow unbalanced panel studies. In this way, a new compliance database for national fiscal rules may exist and be used for future research, similarly to the supranational Compliance Tracker Database in the EU context (Larch and Santacroce [2020]). Both Chapter 2 and 3 introduced the use of Machine Learning Methodology to fiscal rules assessment studies and clear the way for future research on fiscal policy interested in forecasting such as causal inference measurement. These chapters provided evidence that Machine Learning may involve causal estimator that are strong against common econometric bias such as endogeneity and reverse causality biases. Machine Learning also highlighted a high accuracy in the forecasting exercise in Chapter 2, which exceeds the traditional logistic function mainly used in econometrics. In that sense, public policies may benefit from Machine Learning to forecast short and long-term indicators, and prevent macroeconomic imbalances.

Finally, there are several questions to address for the future of fiscal rules. The most general asks if we could avoid the risk of fiscal rules by improving their performance with an appropriate design reform. The second is EU specific and relates to the political challenges of accepting a new reform of fiscal rules (including the SGP) in the EU context. This reform appears necessary to adjust to the post-pandemic economic reality and public finance in EU member countries. It also requires an increasing transparency from government, a greater harmonization of national accounting practices and a comparable accuracy in macroeconomic data measurement between countries. Finally, we also see a main problem which concerns the institutions in charge of monitoring public finance and fiscal discipline: will fiscal watchdogs be sufficient and efficient enough to foster a fiscally sustainable recovery that requires a deactivation of escape clauses, monitoring of fiscal rules and the preservation of fiscal rules credibility which is vital to their performance?

# Bibliography

F. Caselli, L. Eyraud, A. Hodge, F. Diaz Kalan, Y. Kim, V. Lledó, S. Mbaye, A. Popescu, W. H. Reuter, J. Reynaud, E. Ture, and P. Wingender. Second-Generation Fiscal Rules : Balancing Simplicity, Flexibility, and Enforceability - Technical Background Papers . *Staff Discussion Notes No. 18/04*, 2018. 211

M. Larch and S. Santacroce. Numerical compliance with EU fiscal rules: The compliance database of the Secretariat of the European Fiscal Board (Database). 2020. 211, 212

R. Mundell. Capital mobility and stabilization policy under fixed and flexible exchange rate. *Canadian Journal of Economics*, 29:475–485, 1963. 209

C. Wyplosz. Fiscal Rules: Theorical issues and historical experience. *NBER Working Paper No. 17884*, 2012. 210

# List of figures

# List of tables