BETA
Bureau d'économie théorique et appliquée

Documents de travail

«On the Stability and Growth Pact compliance: what is predictable with machine learning? »

Auteurs

Kea Baret, Theophilos Papadimitriou

Document de Travail n° 2019 – 48

Décembre 2019

Bureau d'Économie Théorique et Appliquée
BETA

www.beta-umr7522.fr

@beta_economics

Contact :
jaoulgrammare@beta-cnrs.unistra.fr

On the Stability and Growth Pact compliance: what is predictable with machine learning?

Kea BARET,[1] and Theophilos PAPADIMITRIOU [2]

[1] *Corresponding Author*: BETA CNRS UMR 7522, University of Strasbourg, France. Email: k.baret@unistra.fr

[2] Democritus University of Thrace, Department of Economics, Komotini, Greece

Abstract

The aim of the paper is to propose simplest advanced indicators to prevent internal imbalances in European Union. The paper also highlights that new methods coming from Machine Learning field could be appropriate to forecast fiscal policy outcomes, instead of traditionnal econometrics approaches. The Stability and Growth Pact (SGP) and especially the 3% limit sets on the fiscal balance purpose to coordinate fiscal policies of the European Union member states and ensure debt sustainability. The Macroeconomic Imbalance Procedure (MIP) scoreboard introduced by the European Commission aims to verify the good conduct of public finances. We propose an analysis of the determinants of the SGP compliance by the 28 European Union members between 2006 ans 2018, through a Support Vector Machine model. More than testing if the MIP scoreboard variables really matter to forecast the risk of unsustainability, we also test a set of macroeconomic, monetary, and financial variables and apply a prior feature selection model which highlights the best predictors. We give some proofs that main primary indicators of the MIP scoreboard are not useful for SGP compliance forecast and we propose new variables to forecast the European Union supranational fiscal rule compliance.

Keywords: Fiscal Rules; Stability and Growth Pact, Forecasting, Machine learning.

JEL Codes: E61, H11, H61, H62.

Contents

1 Introduction

The aim of the paper is to propose simplest indicators to forecastinternal imbalances in European Union. The paper also highlights that new methods coming from Machine Learning field could be appropriate to forecast fiscal policy outcomes, instead of traditionnal econometrics approaches.

The excessive imbalances in European member countries motivated us to study the topics of supranational fiscal rules compliance. Many studies focus on the problem of pro-cyclicality possibly generated by the limit of the european 3% rule on the deficit or the 60% on the debt. But we will study the phenomenon of countries not complying with fiscal rules, especially European Union supranational fiscal rule. To promote a sustainable and stable economic area, the Maastricht Treaty set convergence criteria, especially for public finances, with a limit of 3% on the deficit of European Union member states and 60% on their debt. Nevertheless, these limits were set on data from 1992 (year of the implementation of the Maastricht Treaty). This raises the question of the relevance of maintaining such limits today because many countries that have contributed to the formation of the Economic and Monetary Union, such as France or Belgium, exceed the 60% limit on debt.

The Maastricht criteria defined conditions for Economic and Monetary Union (EMU) membership. Subsequently, the European Commission introduced monitoring for macroeconomic state of its members because public finances were degraded. Thus the Stability and Growth Pact was created aiming to coordinate fiscal policies between Member States. The SGP stands out as a supranational fiscal rule in EMU, according to Kopits and Symansky [1998] definition of an ideal fiscal rule.

Recently, Reuter [2019] analyzed the compliance with national fiscal rules through a logistic model. This analysis focuses on numerical fiscal rules in the sense of Kopits and Symansky [1998] , thus excluding Medium-Term Budgetary Frameworks (MTBF). Determinants evaluated in the analysis are mainly rule-related (characteristics that reinforce or strengthen fiscal rules[1]). A potential source of error in this analysis is technical: forecasting whether fiscal rules are complied by using a logistic approach may be inappropriate; the algorithm should forecast different variables in every case. Indeed, since countries have very different national fiscal rules[2] the analysis finally forecasts different events. This is why we focus our analysis on a rule that is identical for all the countries of the sample: the 3% deficit rule. We do not to study the 60% debt rule, since its non-compliance is based on accumulated debts over time and usually this status remain stable for many years (for e.g. France has a debt ratio closed to 100% debt ratio). The SGP reforms introduced new instruments for monitoring but also increased the complexity of the initial simple “3% limit set on deficit”. Since, the European Commission carefully monitors a set of indicators constituting the Macroeconomic Imbalance Procedure (MIP) Scoreboard to prevent the risks of macroeconomic imbalances. Using the most complete series of MIP Scoreboard’s main (and secondary) indicators as well as other macroeconomic variables, we will try to highlight which variables actually explain the fact that countries are derogating from the rule. We also take into account that countries belonging to the Eurozone are more constrained by the 3% limit since a failure in the supranational rule compliance could lead to explicit penalties. We therefore try to answer the following

[1] Registration in the law, level of rigor, degree of public finances coverage, etc.

[2] some set a limit on the structural balance, others on overall balance or the balance excluding public investment...

question: when and why European Union countries do not comply with the 3% deficit limit? We thus confront new predictive methods belonging to the field of Machine Learning whose quality of prediction outperforms the traditional methods used until now. Moreover, our study does not focus only on rules characteristics. We extend the analysis of the determinants of fiscal rules compliance to macroeconomics and financial variables. According to the litterature about "flexible fiscal rules" (Guerguil et al. [2017], Caselli et al. [2018]) we highlight that flexibility is important in fiscal rules design. Even if in Reuter [2019] the Output Gap does not appear important for national fiscal compliance (which seems counter intuitive), we highlight that cyclical events or fiscal space are essential in fiscal rule compliance forecast. This important results reintroduce the debate about the trade-off between credibility and flexibility of fiscal rule. The rest of the paper is structured as follows. Section 2 exposes the data and descriptive statistics, Section 3 describes forecasting methods and Section 4 relates methods that garantee a robust forecast. Finally Section 5 reports the benchmark results and Section 6 concludes the paper.

2 Data and descriptive statistics

2.1 Compliance with the Supranational Fiscal Rule in European Union

With Kopits and Symansky [1998] and their definition of an "Ideal" fiscal rule as a starting point, many papers studied the efficiency of fiscal discipline. Schwengler (2012) defined a fiscal rule as " a sustainable constraint on fiscal policy under the form of a numerical target on a key aggregate of public finances". Nowadays there is a large variety of fiscal rules in OECD countries. Indeed fiscal rules could set a coustrain on Budget Balance, on Debt, on Revenue or Expenditures. These constraints could be at subnational, national or supranational levels. Moreover, fiscal rules could be different because they are characterized by different levels of rigor. The starting point for supranational fiscal rules in the world is the Maastricht Treaty signed in 1992. It marks the launch of the Economic and Monetary Union and defines convergence criteria that must be met by a country to become a member. Two numerical criteria aimed at ensuring the stability of public finances: the debt must be less than 60% of GDP (which corresponds to the average debt of the countries creating EU) and the deficit must be at least 3% of GDP (this corresponds to the level of deficit allowing to stabilize the debt). The Maastricht Treaty drove the construction of the Stability and Growth Pact (SGP).

SGP is not an institutional treaty but an institutional text defined by Regulation No. 1466/97 and Regulation No. 1467/97. At its beginning the SGP resumes only the 3% limit imposed on the deficit. If we want to define today exactly what the supranational rule is, we have to read the texts of 2005, the 6 texts of the 6-pack and the texts of the two-pack. These reforms emerged since some member states deficits had slipped and the criteria on public finances for the entry in the EMU were no longer complied (such as France, e.g Table 1). Behind regulation there are two complementary objectives for the SGP: stability of public finances but also growth.

SGP conduct the monitoring of public finances and has a preventive instrument. Through multilateral monitoring, it requests multiannual programs[3] on public finances for every member state. Non-Member States in the Eurozone must provide stability programs every year, while the member countries of the Eurozone must provide a convergence program. Initially no common deadline was imposed for these programs and monitoring was not as dense as today. This conception resulted in a deficit procedure launched against Portugal (in 2002) and against France and Germany (in 2003) (see Table 1). This led to the first reform of the SGP in 2005. Thus we focus our analysis on the period 2006-2018 to see if the countries have implemented the necessary measures to comply with the supranational rule. The excessive deficit procedure is established by Article 104c of the Maastricht Treaty but the supranational rule is formalized by the Stability and Growth Pact. Now, the conditions are therefore stricter for countries belonging to the euro area since they are subject to explicit sanctions. The deficit should not exceed 3% of GDP, otherwise financial penalties are applied: the European Commission imposes an excessive deficit procedure procedure which lead to a financial sanction between 0.5% and 2% of the GDP. This planned

[3] multiannual programs define what will be the public finances over 3 years

penalty is accompanied by an exclusion clause[4]. In this sense, the SGP stands out as a supranational fiscal rule in the sense of Kopits and Symansky (1998) for the countries of the Monetary Union.

Table 1: **Budget Balance in EUM members from 1999 to 2004 (in % of GDP)**

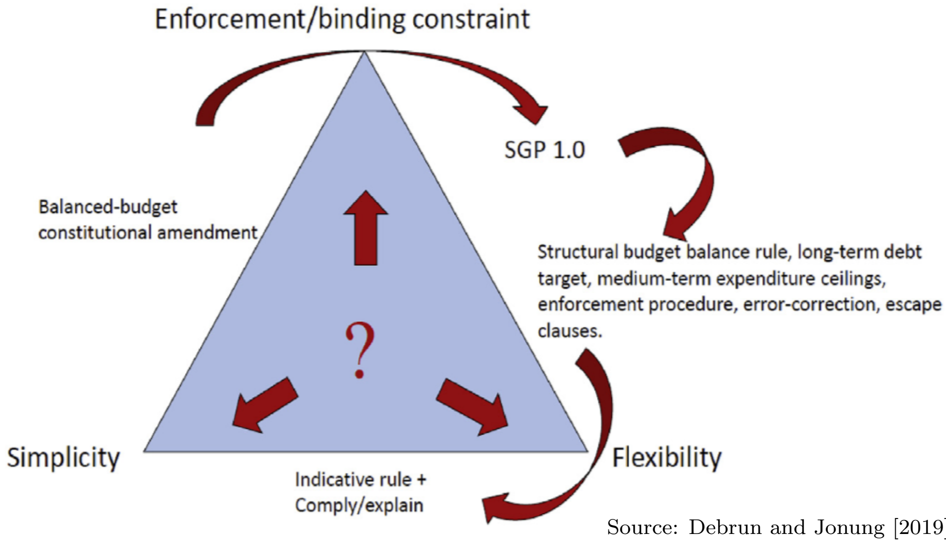
Countries	1999	2000	2001	2002	2003	2004
Austria	-2,4	-1,5	0,3	-0,2	-1,1	-1,3
Germany	-1,5	1,3	-2,8	-3,7	-3,8	-3,7
Belgium	-0,4	0,2	0,6	0,1	0,4	0,1
Spain	-1,2	-0,9	-0,5	-0,3	0,3	-0,3
Finland	2,2	7,1	5,2	4,3	2,5	2,1
France	-1,8	-1,4	-1,5	-3,2	-4,2	-3,7
Greece	-3,4	-4,1	-3,6	-4,1	-5,2	-6,1
Ireland	2,5	4,4	0,9	-0,4	0,2	1,3
Italy	-1,8	-0,6	3,0	-2,6	-2,9	-3,0
Luxembourg	3,4	6,2	6,2	2,3	0,5	-1,1
Netherland	0,7	2,2	-0,1	-1,9	-3,2	-2,5
Danemark	3,3	2,6	3,2	1,7	1,2	2,8
Sweden	2,5	5,0	2,5	-0,3	0,2	1,4
United Kingdom	1,0	3,8	0,7	-1,7	-3,4	-3,2

Source: Eurostat

Kopits and Symansky [1998] defined the Ideal fiscal rule as a fiscal rule well- defined for the compelling, transparent, simple, flexible, consistent, enforceable indicator, accompanied by a fiscal framework and related reforms. As there is the monetary policy trilemma (Mundell [1963]), Debrun and Jonung [2019] propose fiscal policy trilemma in the design of fiscal rules:

[4]Indeed, there are exceptions to the rule when the country is situation of "exceptional" circumstance namely when it undergoes a recession of -2% growth rate. However between -0.75 and -2% it is possible that an exceptional circumstance is recognized.

Figure 1: **The fiscal rules design trilemma**



Source: Debrun and Jonung [2019]

What they called "SGP 1.0" corresponds to the initial and simple version of SGP. The several reforms re-introduced Debt criterion but also additionnal objectives with the aim to answer to the critics that had been adressed to him. Even if Villieu [2003] exposed that SGP could be a second-best optimum as solution to coordinate fiscal policies in monetary union, Creel [2003] pointed out the risk that countries could deviate from the solution. The solution now appears as under-optimal because the supranational rule is not credible enough. Creel [2003] promotes the use of a golden rule instead of the SGP which appear to strict et is not efficient again "fiscal short-termism" (Bonatti and Cristini [2008]). The SGP several reforms finally contributed to the growing complexity of the european fiscal framework: with the long-term debt rule new objectives were set as structural budget balance rule (the so-called "medium-term objective"), medium-term expenditure ceilings, escape clauses and enforcement procedure which goes beyond Article 104c of the Maastricht Treaty concerning the excessive deficit procedure. This context leads to the fiscal rules design trilemma (view Figure 1 from Debrun and Jonung [2019]). The MIP scoreboard was introduced to prevent the risks of internal and external imbalances and also monitor employment indicators. Internal imbalances must be avoid by fiscal policy which must be constrained by fiscal rules (at several levels). It appears really difficult for decision makers to present all these indicators as "green" and also comply with all their fiscal rules. Indeed, since the 1990s the number of national fiscal rules increased in European Union with the aim to promote sound public finances. This trend also introduce complexity for fiscal policy, decreasing simplicity and clarty of the objectives. Our first motivation here is to clarify the use of the MIP Scoreboard to forecast internal imbalance. If the MIP scoreboard is usefull, its indicators must help to forecast is the countries comply with the supranational (in other terms, if the discretionary fiscal policy is constrained). In all the rest of our study we will just be interested in the simple "SGP 1.0" compliance. We want to see if the most simply rule included in the large SGP supranational rule could be comply with.

2.2 Potential determinants of fiscal rules compliance

Novadays there is a large litterature on fiscal rules. Some study the disciplining effect of national and sub-national fiscal rules. Through different methods as IV (Foremny [2014]), LSDV (Reuter [2015]), system-GMM (Bergman et al. [2016]), or propensity-score Matching (Tapsoba [2012] and Barbier-Gauchard et al. [2019]), a common result pointed out is that countries with fiscal rules present, on average, better fiscal discipline compared to countries without fiscal rules. Finally implementation of fiscal rules is correlated with better fiscal performance (IMF [2009], Barbier-Gauchard et al. [2019]). Some studies focused on European Union case as Debrun et al. [2008] since other mix devlopping and developped countries as Combes et al. [2018]. An other part of this litterature focuses on macro-stabilizing power of fiscal rules. Finally, Sacchi and Salotti [2015] highlighted that national fiscal rules helped in GDP stabilization and Guerguil et al. [2017] showed that flexible budget balance rules helped in expentidures stabilization (for standard definition of flexible rules e.g. Schick [2010], Dabán [2011] or Caselli et al. [2018]). The literature about the supranational rule in European Union essentially focused on is legitimity. SGP is here to promote fiscal policies between Eurozone members (Villieu [2003]) but this rule does not look credible enough and in presence of externalities between members countries could deviate from (Creel [2002]). Finally a high debate is around a golden rule to replace SGP: a such rule could reduce countries arbitrage between supranational fiscal rule compliance and growth objectives (Mathieu [2003] or Creel [2003], Creel et al. [2007]).

2.3 Average compliance and public finances statistics

Table 2: **Public Finances Statistics in European Countries between 2004 and 2018**

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

Table 2 depicts an overview of public finances and supranational fiscal rules statistics. We observe that SGP was respected 63,70% on average. France is the country that don't comply with the rule most times. We also observe a large heterogeneity between highest and lowest values in public finances of the different EU member states. As explained in Eyraud et al. [2018] we observe poor compliance with fiscal rules in general. Eyraud et al. [2018] highlights a "magnet-effect" for countries submitted to a 3% fiscal rule, showing that countries make effort to comply with the rule. Reuter [2015] also pointed out that governments make efforts to move closer to their national fiscal rules limit when there are not complied. Solutions as raising reputional costs for non compliers and benefits for compliers are proposed in Eyraud et al. [2018]. But, government still take care of fiscal rule compliance because fiscal rules are still effective even if countries don't comply with. Indeed there are direct sanctions for non-compliers provide by SGP. Moreover, non-compliers are also submitted to quasi-automatic mecanisms as "discriminating principle"(Buiter and Kletzer [1991]) through financial markets, partners countries or European Central Bank. Some also argues that ricardian effect could be a sanction (which is questionned, e.g. Weil [1987], Burbidge [1983], Jaeger [1998]).

We are so interested to observe what are the determinants that really matters to predict SGP compliance. It is also a way to study the relevance of primary monitoring indicators used by European Commission to garantee internal balance.

Table 3 describes out data set. We retain all the 28 European Union members and we study the period 2006-2018. Y is our binary dependent variable traducing if a country comply with the SGP (3% limit) in year t. All potential features are lagged values because we are in a forecasting model. We test fkr 44 potential features over 3 lags. Among these 44 potential features we have main primary and secondary indicators of MIP scoreboard (X8 to X38). We look at Countries characteristics as potential feature (X1 to X5). But also we look at if the Eurozone members comply with the rule more often than non-Eurozone members (X6) and if it's the formal procedure provide by SGP made a difference (X7). We test for macroeconomics variables that are not monitored by European Commission for internal imbalance (X39 to X44): Oil Prices, Bonds yield, foreign currency long-term sovereign debt ratings[5], fiscal space simply measured as the difference between a country level of debt and median of European Union (e.g. Cheng and Pitterle [2018] for an overview of fiscal space definitions), and indicators of macroeconomics cycle (presence of crisis and Outputgap measured by production function approach). In Reuter [2019] output gap did not appear as an important determinant for national fiscal rule compliance. But SGP provide escape clauses and members such as UK announced temporarily abandoned their fiscal rule during the Global Financial Crisis. We so expect that crises and cyclical fluctuations have an impact for fiscal rules compliance forecasting.

Table 3: **Variables Overview**

Variables	Correspondance Variables	Study Period	Countries considered
Y	Dummy variable =1 if 3% limit was complied in t	2006	Austria
X1	Dummy variable traducing if the country was an advanced country in t-p	2007	Belgium
X2	Dummy variable traducing if the country was an Emerging country in t-p	2008	Bulgaria
X3	Dummy variable traducing if the country was a Ressource-rich country in t-p	2009	Croatia
X4	Dummy variable traducing if the country was an EU membership in t-p	2010	Cyprus
X5	Dummy variable traducing if the country was a Federal Country in t-p	2011	Czech Republic
X6	Dummy variable traducing if the country was a Eurozone member in t-p	2012	Denmark
X7	Dummy variable traducing if the country was submitted to an enforcement	2013	Estonia
	procedure related to the supranational fiscal rules in t-p	2014	Finland
X8	Gross domestic product, deflator, in t-p	2015	France
X9	Total investment in t-p	2016	Germany
X10	Gross national savings in t-p	2017	Greece
X11	Inflation, average consumer prices, in t-p	2018	Hungary
X12	Population in t-p		Ireland
X13	General government revenue in t-p		Italy
X14	General government total expenditure in t-p		Latvia
X15	General government net lending/borrowing in t-p		Lithuania
X16	General government gross debt in t-p		Luxembourg
X17	Net External Positions in t-p		Malta
X18	Current account balance in t-p		Netherlands
X19	Current account balance variations over 3 years in t-p		Poland
X20	Real Effective Exchange Rate in t-p		Portugal
X21	Global export market share -% change over 5 years - in t-p		Romania
X22	Nominal unit wage cost -% change over 3 years - in t-p		Slovak Republic
X23	Debt of private sector in t-p, consolidated -% of GDP		Slovenia
X24	Liabilities of the financial corporations sector, -% change over 1 year - in t-p		Spain
X25	Unemployment rate - 3-year average - in t-p		Sweden
X26	Unemployment rate in t-p		United Kingdom
X27	Gross domestic product (real GDP) -% change over 1 year - in t-p		
X28	Gross fixed capital formation in t-p -% of GDP -		
X29	Gross domestic expenditure on R & D in t-p -% of GDP -		
X30	Direct investment in the reporting economy (flow) in t-p -% of GDP -		
X31	Direct investment in the reporting economy (stocks) -% of GDP		
X32	Net trade balance of energy products in t-p -% of GDP -		
X33	Real effective exchange rate, euro area trading partners -% change over 3 years		
X34	Terms of trade (goods and services) -% change over 5 years - in t-p		
X35	Market share of world exports, volumes -% change over 1 year - in t-p		
X36	Labor productivity -% change over 1 year - in t-p		
X37	Residential construction in t-p -% of GDP -		
X38	Employment -% change over 1 year - in t-p		
X39	Dummy variable traducing if there is a Crisis in t-p		
X40	Output gap (production function approach) in t-p		
X41	Oil Prices in t-p		
X42	Bonds yield in t-p		
X43	Foreign currency long-term sovereign debt ratings, index from 1-21 , in t-p		
X44	Fiscal Space in t-p		

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 test for $p = 1, 2, 3$ for each feature. There are 44 variables included in 3 lagged so 132 features tested. Fiscal Space is measured as the difference between country public debt and EU median debt for each year.

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

3 Forecasting algorithms

3.1 The logistic function

Reuter [2019] used logistic function to find the main determinants for compliance with 51 national fiscal rules in European Union countries. He pointed out that rules-characteristics seem to be important for compliance whereas the lagged value of the output gap is not. To implement forecasting approach we can't work on differents fiscal rules. National fiscal rules could set a limit on expenditure, budget balance, debt or revenues of government. Even if the limit is set on the same aggregate, it could differ from one country to another [6]. Because we can't ask to the algorithm to forecast different events we will only focus on the "3 % limit on deficit" wich is the same for all European Union members. Ince and Trafalis [2006] compared Support Vector Machine (SVM) to logit models in forecasting EUR/USD exhange rate. They finally highlithed that SVM outperforms logit model. Plakandaras et al. [2013] trained an SVM model to forecast EUR/USD exchange rate directional changes, and SVM also outperformed other machine learning methods. Similarly, we conducted tests using both a traditional econometric method (Logit) and an emerging methodology in Economics from the Machine Learning field (Support Vector Machine). The aim of this approach is to find the best forecasting model.

Our issue leads to a simple classification problem: we must predict whether the countries in our dataset will comply with the rule (class 1) or will not comply with the rule (class 0). Since we only have 2 classes, this is a binary classification problem to solve $Y \in \{0;1\}$. The goal of the classification is to find a linear or non-linear separator to separate the two classes.

Considering our data, Inputs $\mathbf{x}_i$ corresponds to feature-vectors including p potential features ($p = 1,\dots,k$) and $i = 1,\dots,n$.
Our input matrix X contains $N = 364$ number of inputs (data points) each contains $p = 132$ number of features. Y is our binary outcome taking the value 0 or 1 . We can illustrate our data as below:

$$X = \begin{bmatrix} \mathbf{x}_{11} & \mathbf{x}_{12} & . & . & . & \mathbf{x}_{1p} \\ \mathbf{x}_{21} & . & & & & . \\ . & & . & & & . \\ . & & & . & & . \\ . & & & & . & . \\ \mathbf{x}_{n1} & & . & & & \mathbf{x}_{np} \end{bmatrix}_{n\times p}, Y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix}_{n\times 1}$$

The well-known logistic function constrains Y in a range of (0,1) and uses the sigmoid function :

$$p(Y|X) = \frac{exp^{\beta_0+\beta_1 X}}{1+exp^{\beta_0+\beta_1 X}} \Longleftrightarrow log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X \tag{1}$$

Which is equivalent to

$$p(y_i = 1) = \pi_i = \frac{exp^{\mathbf{x}_i\beta}}{1+exp^{\mathbf{x}_i\beta}} \tag{2}$$

where $\mathbf{x}_i$ is the $i-th$ row of an matrix of n observations with p predictors and a column of ones to accommodate the intercept, and β is the column vector of the regression coefficients.

[6]for exemple Denmark and United Kingdom had a Budget balance rule in 2005 but for United Kingdom it was a Golden rule whereas Denmark setted that structural general government surpluses should be around 2 % of GDP

The aim is to find values of β_0 and β_1 (or β for simplification instances) which conducts to $p(Y|X)$ and most accurately classifies observed data points. This problem is equivalent to maximise the product of these probabilities, so-called the likelihood:

$$l(\beta) = \sum_{i=1}^{n}[y_i log(\pi_i) + (1-y_i)log(1-\pi_i)] = \sum_{i=1}^{n}[\ y_i\ log(\frac{\pi_i}{1-\pi_i}) + log(1-\pi_i)] \tag{3}$$
$$= \sum_{i=1}^{n}[y_i \mathbf{x}_i \beta - log(1+exp^{\mathbf{x}_i \beta})]$$

3.2 The Support Vector Machine (SVM)

3.2.1 The Support Vector Machine in linearly separable case

Machine learning techniques are increasingly used for forecasting in economics, especially in macro-finance. SVM showed high forecasting performance in several previous studies. Gogas et al. [2015] were interested in yield curve's ability to forecast economic activity. Through models for forecasting the positive and negative deviations of real US GDP from its long-run trend over the period from 1976 to 2014, they compared a traditionnal econometric approach (probit) to SVM for the forecast, showing that SVM outperforms. Gogas et al. [2018] used SVM in Forecasting Bank Failures and obtained an incredible 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) and the most accurate model is linear SVM.

For simplification instance we still use only $\mathbf{x}_i$ in the rest of the study. SVM is a supervised machine learning method[7]. SVM aims to identify a small set of data points from the initial dataset, called Support Vectors that define the position of the linear separator between the two classes. Considering our data, y_i is our 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 boundary could be define as:

$$f(\mathbf{x}_i) = \mathbf{w}^T\mathbf{x}_i - b = 0, \qquad y_i f(\ \mathbf{x}_i) > 0 \tag{4}$$

with $\mathbf{x}_i\ \in R^2$ and $i = (1, ..., n)$; $\mathbf{w}$ is the weight vector, $\mathbf{b}$ is the bias. In that sense all data satisfy:

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

This optimal separator is defined as the decision boundary that classifies each data vector to the correct class and has the maximum distance from each class. This distance is often called "margin" and exactly correspond tot the distance of the hyperplane with each class.

[7]Supervised learning is the concept behind applications as facial recognition in smartphones for example. Technically, given a set of data, described by a set of characteristics X (features), a supervised learning algorithm will find a mapping function which describes a relationship between X and Y and so-called "forecasting model".

Figure 2: **Hyperplane and Support Vectors**

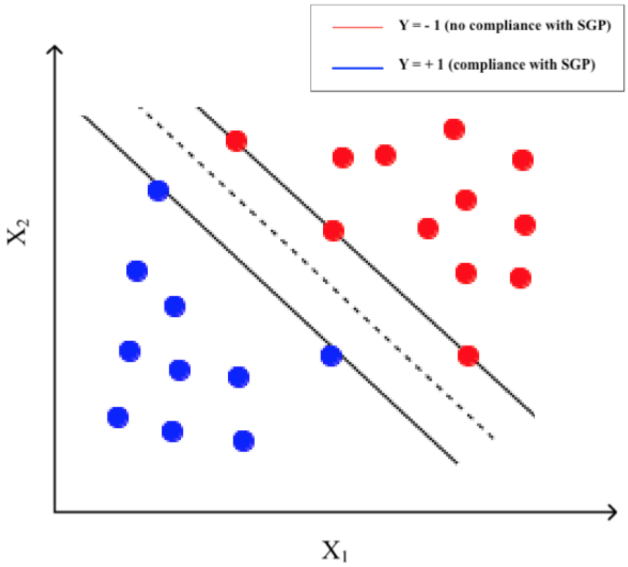
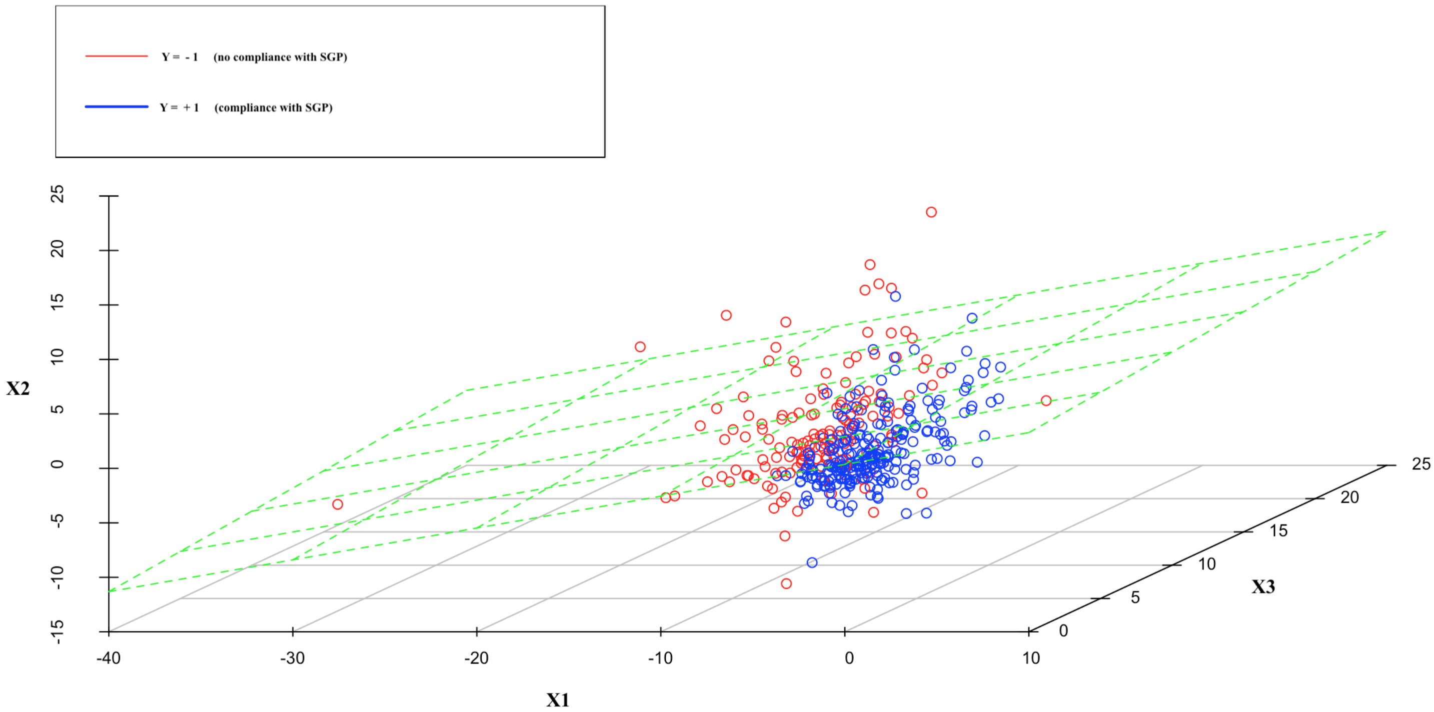


Figure 3: **Research of optimal hyperplane through a simple example representation**



In Figure 2 and Figure3 red points correspond to observations which do not comply with the 3% limit on deficit of the SGP, and blue point are observations which comply with. For a simple comprehension in 2 dimensions, Figure 2 shows the hyperplane which corresponds to dashed line whereas the margin lines are represented by the continuous lines. Support Vectors correspond to the circles identified on the margin lines. Figure 3 extends example to 3 dimensions where the green hyperplan should be adjust to classify each observations correctly and being equal to the optimal hyperplane.

Issue of finding the hyperplane could be solved by using the Lagrange relaxation in a quadratic problem:

$$\min_{\mathbf{w},b} \max_{a} \left(\frac{1}{2}\mathbf{w}^2 - \sum_{i=1}^{N} a_i [y_i(\mathbf{w}^T \mathbf{x}_i - b) - 1] \right) \tag{5}$$

In Equation (5) $a = [a_1, .., a_n]$ correspond to non-negative Lagrange multipliers. This Equation (5) is not used to estimate solution, instead we use a dual problem described by:

$$\max_{a} \left\{ \sum_{i=1}^{N} a_i - \sum_{j=1}^{N} \sum_{k=1}^{N} a_j a_k y_j y_k \ \mathbf{x}_j^T \mathbf{x}_k \right\} \tag{6}$$

with $\sum_{i=1}^{N} a_i yi = 0$ and $0 \leq a_i, \forall_i$. By solving (6) we obtain the location of the hyperplane given by:

$$\widehat{\mathbf{w}} = \sum_{i=1}^{N} a_i y_i \mathbf{x}_i \tag{7}$$

$$\widehat{b} = \widehat{\mathbf{w}}^T \mathbf{x}_i - y_i, i \in V, \tag{8}$$

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

But in reality, data sets contain some noise and outliers. To take into account of this potential bias we use **Error-tolerant SVM** proposed by Cortes and Vapnik [1995] who introduced non-negative slack variables $\xi_i \geq 0, \forall i$ and a parameter C. Finally the problem to solve is:

$$\min_{\mathbf{w}b\xi} \max_{a\mu} \left\{ \frac{1}{2}\mathbf{w}^2 + C\sum_{i=1}^{N} \xi_i - \sum_{j=1}^{N} a_j [y_j(\mathbf{w}^T \mathbf{x}_j - b) - 1 + \xi_j] - \sum_{k=1}^{N} \mu_k \xi_k \right\} \tag{9}$$

where the non-negative slack ξ_i correspond to the distance of vector $\mathbf{x}_i$ from the hyperplane when classified erroneously. The optimal hyperplane is finally given by:

$$\widehat{\mathbf{w}} = \sum_{i=1}^{N} a_i y_i \mathbf{x}_i \tag{10}$$

$$\widehat{b} = \widehat{\mathbf{w}}^T \mathbf{x}_i - y_i, i \in V, \tag{11}$$

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

3.2.2 The Support Vector Machine in non linearly separable case

It's common to observe that the two classes are not linearly separable (as it was in Figure 2). This corresponds to the left case of Figure 4. To make the two classes linearly separable SVM is coupled with kernel 'tricks'. Kernel help to compute dot product of two vectors in some (high dimensional) feature space[8] (right case in Figure 4). This method uses the projection apparoach while ensuring minimum computational cost. Data set is so projected in an inner product space. The projection uses the dot products within the original space through different kernel functions. By introducing kernel projection in the solution to the dual problem in Equation (5) gives:

$$\max_{a} = \sum_{i=1}^{N} a_i - \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{N} a_j a_k y_j y_k K(\mathbf{x}_j, \mathbf{x}_k). \tag{12}$$

[8]kernel functions are so called "generalized dot product"

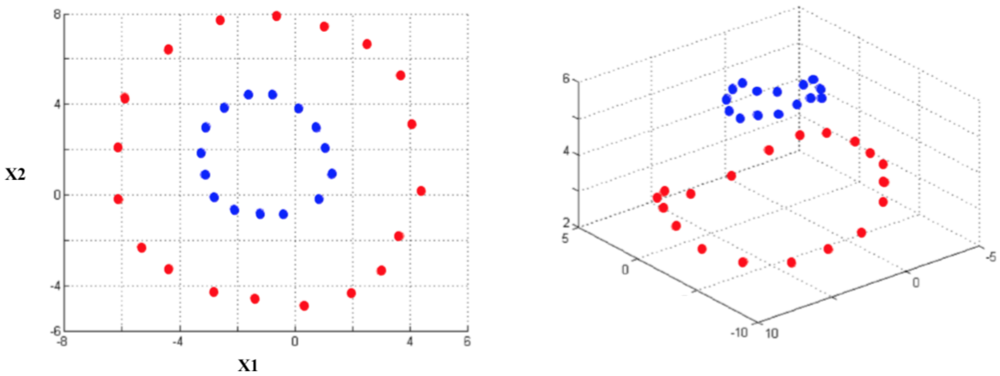
with $\sum_{i=1}^{N} a_i yi = 0$ and $0 \leq a_i \leq C, \forall_i . K(\mathbf{x}_j \mathbf{x}_k)$ corresponds to one of the following kernel function:

$$\textit{Linear} \qquad K_1(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j + r, \tag{13}$$

$$\textit{RBF} \qquad K_2(\mathbf{x}_i \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2}, \tag{14}$$

$$\textit{Polynomial} \qquad K_3(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j + r)^d, \tag{15}$$

Figure 4: **Kernel projection to make the two classes linearly separable**



4 For well implementation of forecasting model

4.1 Feature Selection: The logistic LASSO

Friedman et al. [2009] proposed LASSO as a regularization alternative that overcomes the disadvantage of ridge regression inability of reducing the number of predictors in the final model. Pereira et al. [2016] used logistic LASSO in predicting corporate failure.

The logistic function can be used to answer a problem of classification but it could also be found in the method of logistic LASSO regression which operates preliminary feature selection in the case of a binary dependent variable.

LASSO performs regularization and feature selection. By definition, the feature selection reduces number of explanatory variables. LASSO applies a regularization process where coefficients of some variables are penalized and shrinked to zero. The main goal is to minimize the prediction error and other main advantages of feature selection are the following:

- removing variables that that are not important for the forecast or that are redundant;
- to reduce the size of the problem algorithms have to solve and reduce overfitting.

Finally, the LASSO estimator applied in logistic regression is:

$$\hat{\beta}(\lambda) = \underset{\beta}{\operatorname{argmin}}(n^{-1} \sum_{i=1}^{n} \rho_{(\beta)}(X_i, Y_i) + \lambda ||\beta||_1) \tag{16}$$

Parameter λ is found by grid research and used the one-standard error rule[9]. Finally we will choose the simplest model whose accuracy is similar to the best model.

[9]The approach of one-standard error is heuristic because one-standard error typically is not large relative to the range of λ values.

4.2 Cross-Validation for Robustness check:

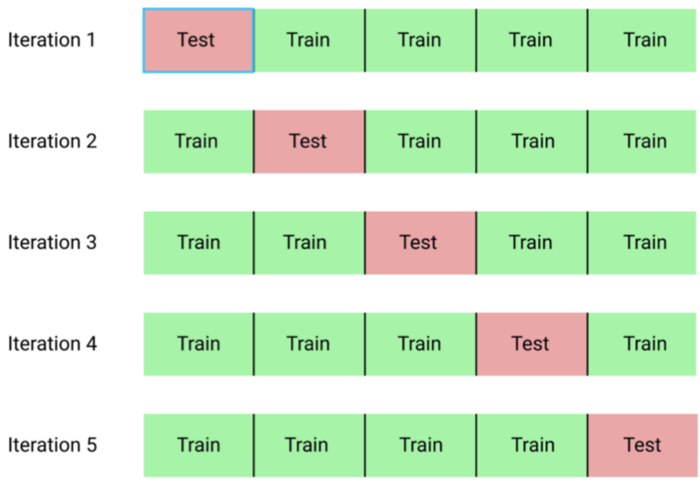
There is no potential reverse causality bias when using a machine-learning scheme whereas reverse causality is important to take into account in many traditionnal econometrics approach. But, an important bias to take into account here is overfitting. Overfitting traduces that the model produced is affected by possible noise or a possible short-run dynamics. Overfitting is traduced by high performance at the training step whereas accuracy significantly drops at the testing step.

Overfitting can be avoided using cross-validation. We will use 2 cross-validation approach: hold-out cross validation and k-fold cross validation.

Hold-out cross validation consist in spliting up our dataset into a 'train' set and 'test' set. The model is trained on training set and the test set is used to check the performance of the model on an unknown data. A common split is to use 80% of data for training set and remain 20% of the data for testing set.

k-fold cross validation repeats hold-out k times. Indeed, our data set is splited 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 constitues the 'train' set. The average performance from every fold is used to obtain evaluate final model.

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



4.3 Measures of forecasting performance

Forecasting performance: The accuracy is a common tool to measure forecasting performance and is defined as the ratio of total correctly predicted observations against all observations.

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{17}$$

TP: the number of True Positive instances

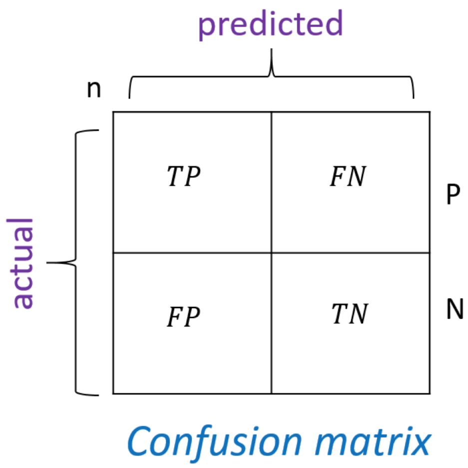
TN: the number of True Negative instances

FP: the number of False Positive instances (observations that were predicted as they complied with the rule but they not complied in reality)

FN: the number of False Negative instances (observations that were predicted as they not complied with the rule but they complied in reality)

Confusion Matrix In confusion matrix we can see all the cases described before: True Positive instances, True Negative instances, False Positive instances, False Negative instances. We need the model that produces the least false positives instances. Indeed a high number of False Positive observations would traduce a risk for unsustainability if we forecast that a country will comply with SGP and in reality it will not. If there is no suspicion that the country will not comply with, no strong recommendations will be prepared by the Commission since it is not expected to worsen public finances.

Figure 6: **A simple representation of confusion matrix**



5 Results

5.1 Preliminar testing results

We performed various set of tests. First, we included the main indicators of the MIP scoreboard to test if monitoring public finances through these variables is efficient to promote fiscal sustainability. We compared different SVM models (described in Equations (13), (14), (15)) and logistic regression. Finally, the kernel linear SVM outperforms other kernel models. The first and second columns of table 4 show that the accuracy obtained with the best SVM model (linear SVM) and logistic regression is not really satisfying with at best 68,7 % for the linear SVM model and 67,0% for the logistic model. In columns 3 and 4, we include the complete set of variables, to test the hypothesis that other variables can forecast better the risk of internal imbalances than MIP scoreboard indicators. We also test different SVM models and compare them to the logistic model. The idea is to look at what is the model which could perform even if there is no optimisation for the number of feature. After, the identification of this best model, we will work only with this one. In columns 3 and 4 of table 4 we observed that the accuracy increase for both models and even more for the linear SVM with an accuracy of 83,2%. We can therefore deduct that there are features in all our data set that really help to monitor internal fiscal imbalances. Also, we will retain linear SVM for all the rest of our predictions as it appears to be the model which outperforms.

Table 4: **"Compliance with 3% limit" forecasting accuracy: models comparison (%)**

Model	MIP scoreboard main indicators included	MIP scoreboard main indicators included	All features includes	All features includes (132)
Linear SVM model	59,7	68,7	77,8	83,2
Logistic model	61,1	67,0	76,4	76,1
Validation method	Hold-out	k-Fold Cross Validation	Hold-out	k-Fold Cross Validation

Note: Hold-out splits up dataset into a 'trainset' (80%) and 'testset' (20%). Results are on testset. k-Fold Cross Validation is a 5-Fold Cross Validation et gives mean results. Parameter C in SVM is equal to 2^2 and obtained using power of 2 grid search.

5.2 Feature Selection and final predictive model

The logistic LASSO was implemented following the results presented in table 4. We included features retained by logistic LASSO and tested all the combinaisons of theses features in kernel linear SVM. We finally worked with the combinaison which produced the best accuracy. It leads to 7 main variables, some considered over several lags:

Table 5: **Best predictors:**

General government fiscal balance in t-1
Liabilities of the financial corporations sector, % change over 1 year, in t-1
Dummy variable traducing if there was a crisis in t-1 and t-2
The Output Gap in t-1
The oil Prices in t-1
Bond yield in t-1 and t-2
Fiscal space in t-1, t-2 and t-3

table 5 depicts the best predictors retained and that could be interpreted as advanced indicators for SGP compliance. As we could expected the Lag-1 of General government fiscal balance is a key feature. Our dependent variable is defined by a limit set on overall fiscal balance. A degraded fiscal balance in t-1 will have an impact on the current value of fiscal balance, but governments can adjust their finances after this period. Among the lessons we learned from the Global Financial Crisis, we find that the solvency and quality of the commercial banks' liabilities has a direct impact on public finances sustainability since these same banks also hold treasury bills. We also observe that shocks as big crises could lead countries to do not comply with the rule because of their need of public investments. Crises in t-2 and t-1 are important, not if it arrived in t-3, showing that escape clause should be adapted to crises lenght and should not only focused on a percent of GDP recession. The key variable to measure the deviation is not an arbitrary value of "exceptional" circumstance traduces by a recession of -2% growth rate (knowing that between -0.75 and -2% European Commission could consider an exceptional circumstance). Output gap from AMECO Database could be use as an indicator of position in the cycle. Lag-1 of Oil Prices is also a key variable for the forecast. We made experience of Oil prices shocks in 1973 and 1979, but also at the begining of 2008 there was again a surprise increase in oil prices. Moreover since november 2018, Yellow vests movement in France highlights population reaction in front of oil prices increase. Measures of flexibility for governments, especially in crises periods, as fiscal space or bond yields appear as really matter for the forecasr. Our result is part of the continuity of Romer and Romer [2018] study which highlighted the importance of fiscal space during financial crises and normal recessions.

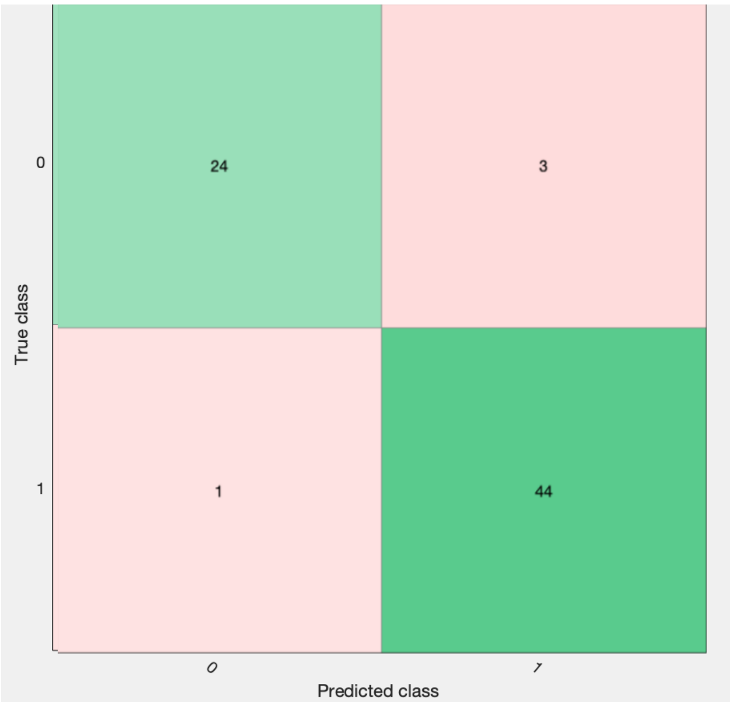
Table 6: **"Compliance with 3% limit" forecasting accuracy with linear SVM(%)**

Model	Features selected by LASSO included	Features selected by LASSO included
Linear SVM model	90.7	94.4
Cross Validation Method	k-Fold Cross Validation	Hold-out validation

Note: Hold-out splits up dataset into a ‘trainset’ (80%) and ‘testset’ (20%). Results are on testset. k-Fold Cross Validation is a 5-Fold Cross Validation et gives mean results. Parameter C in SVM is equal to 2^2 and obtained using power of 2 grid search.

Finally we obtain a satisfying accuracy with our linear SVM model: 90.7 % with 5-fold Cross Validation and 94.4% with Hold-out Cross validation. So we can affirm with more than 90% of precision that a country will fail in SGP compliance. Main indicators of MIP scoreboard that should help to monitor internal imbalances present only 68.7% precision at best (and only 59.7% with hold-out cross validation) so these indicators do not appear as usefull to monitor internal imbalances. The complexity of these indicators, often defined over several years, finally makes the European Union fiscal framework difficult to monitor with them.

Figure 7: **Linear SVM confusion matrix (Hold-out Cross Validation)**



The confusion matrix shows that only 3 observations of the 'test' set are False positive. It is also a satisfying results because we increase the forecasting power obtain in confusion matrix of an SVM using MIP scoreboard variables (view Appendix 2). Also we have only 1 False negative. As a result, the European Commission will not often have to make detailed recommendations to avoid a risk that in fact does not exist.

6 Conclusion

Our study proposes a new forecasting approach in the issue of fiscal rules compliance. SVM model outperforms a standard logistic regression often used in economics study. It feeds the debate about "The Impact of Machine Learning on Economics" (Athey [2018]). We highlight a set of simple indicators to forecast SGP compliance. Our main policy implication is that MIP scoreboard indicators (first and secondary indcators) are not efficient in internal imbalances forecast, except financial sector liabilities. However, Debrun et al. [2019] have highlighted the importance of great monitoring by institutions and especially the provision of unbiased quantitative analysis. MIP scoreboard indicators could be used in European Commission recommandations to help countries with their fiscal difficulties but not to implement excessive imbalances/deficit procedures. With the aim to simplify the european fiscal framework and to lighten the fiscal rule design trilemma of Debrun and Jonung [2019], the use of simple advanced indicators to prevent SGP deviations could be a first step in the European Fiscal framework reform which is needed today. Also, the forecasting performance domination of SVM/machine learning approach on forecastinf performance of traditionnal econometrics apparoach (as logit model) points out the advantage of the the use of machine learning in Economy. Our study can therefore open the way to the use of these models in other macroeconomics studies interested in fiscal policies outcomes. The analysis could also be transposed to national fiscal policies outcomes forecasting with large available dataset at national level.

References

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

A. Barbier-Gauchard, K. Baret, and A. Minea. National Fiscal Rules Adoption and Fiscal Discipline in the European Union . BETA Working Paper No. 2019-40, 2019.

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.

L. Bonatti and A. Cristini. Breaking the Stability Pact : Was it predictable ? Journal of policy modeling, 30:793–810, 2008.

W.H. Buiter and K. Kletzer. Reflections on the Fiscal Implications of a Common Currency. in A. GIOVANNINI and C. MAYER (eds.), European Financial Integration, CEPR, Cambridge University Press, pages 221–244, 1991.

J.B. Burbidge. Government Debt in an Overlapping-Generations Model with Bequests and Gifts. The American Economic Review, 73:222–227, 1983.

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.

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

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.

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

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

J. Creel. Ranking Fiscal Policy Rules: the Golden Rule of Public Finance vs. the Stability and Growth Pact. 2003.

J. Creel, P. Monperrus-Véroni, and F. Saraceno. Has the Golden Rule of Public Finance Made a Difference in the UK ? OFCE Working Paper No.2007-13, 2007.

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

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.

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.

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.

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.

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.

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

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

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.

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

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.

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.

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

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

K. Jaeger. An Overlapping-Generations Model with Heterogeneous Bequest and Gift Motives. Trade, Growth, and Economic Policy in Open Economies. Springer, pages 139–163, 1998.

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

C. Mathieu. La politique économique : la clé anglaise. Lettre de l'OFCE, 233, 2003.

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

J.M. Pereira, M. Basto, and A. Ferreira da Silva. The Logistic Lasso and Ridge Regression in Predicting Corporate Failure. Procedia Economics and Finance, 39:634–641, 2016.

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.

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

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

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.

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.

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

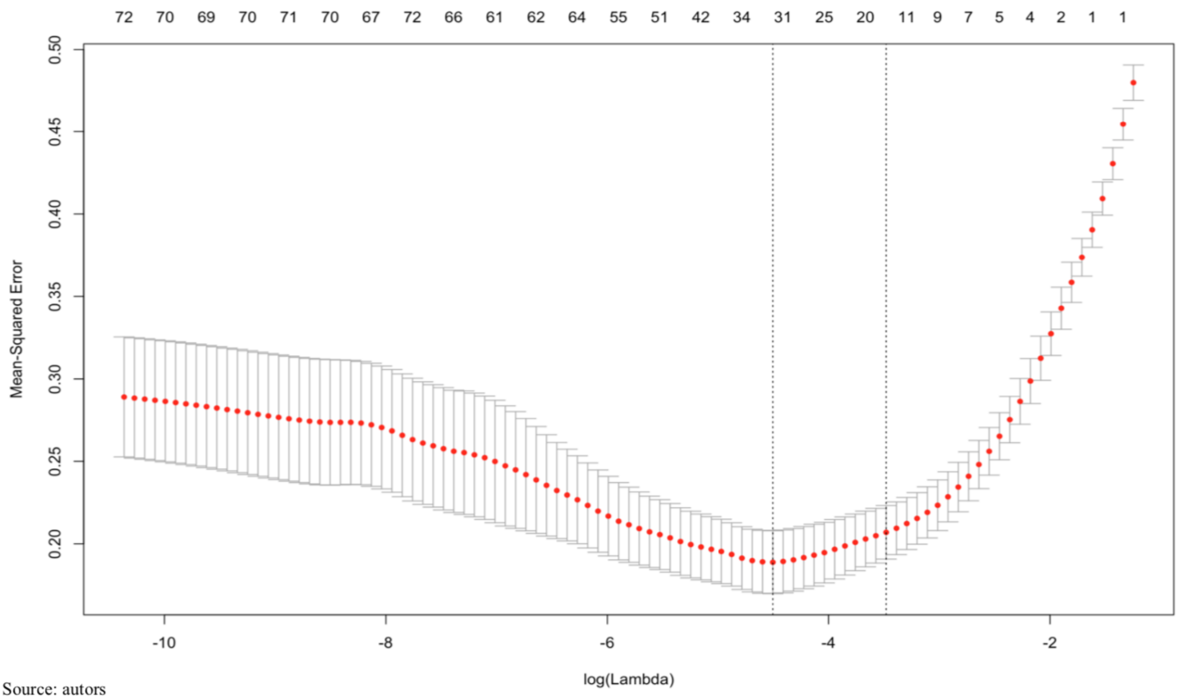
R. Tapsoba. Do National Numerical Fiscal Rules really shape fiscal behaviours in developing countries? A treatment effect evaluation. Economic Modelling, 29:1356–1369, 2012.

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

P. Weil. Love thy children: Reflections on the Barro debt neutrality theorem. Journal of Monetary Economics, 19:377–391, 1987.

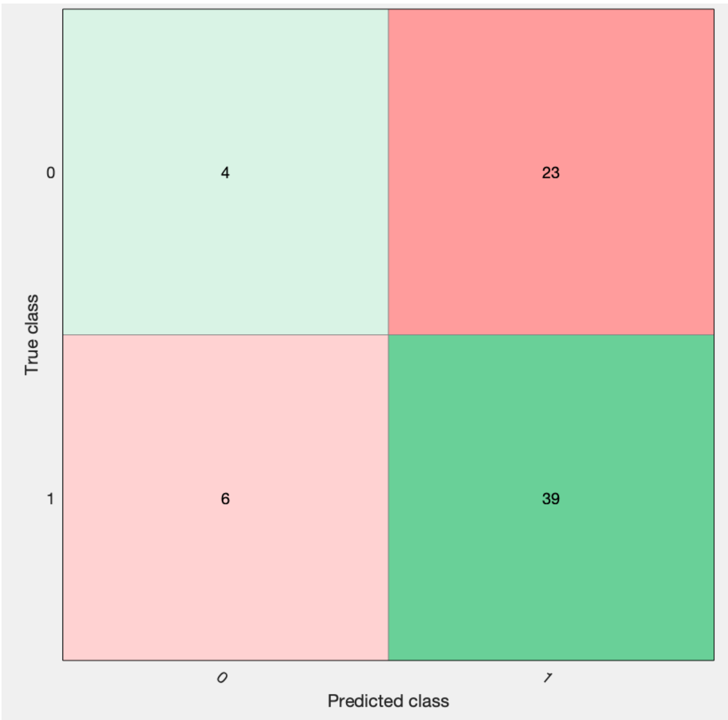
Appendices

Appendix 1 : LASSO results



Source: autors

Appendix 2 : Confusion Matrix from Linear SVM with primary indicators from MIP Scoreboard



Appendix 3. Variables sources

Variables	Correspondance Variables	Source/Database
Y	Dummy variable =1 if 3% limit was complied in t	Authors' calculations
X1	Dummy variable traducing if the country was an advanced country in t-p	IMF Fiscal Rules Database
X2	Dummy variable traducing if the country was an Emerging country in t-p	IMF Fiscal Rules Database
X3	Dummy variable traducing if the country was a Ressource-rich country in t-p	IMF Fiscal Rules Database
X4	Dummy variable traducing if the country was an EU membership in t-p	IMF Fiscal Rules Database
X5	Dummy variable traducing if the country was a Federal Country in t-p	IMF Fiscal Rules Database
X6	Dummy variable traducing if the country was a Eurozone member in t-p	IMF Fiscal Rules Database
X7	Dummy variable traducing 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
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 traducing if there is a Crisis in t-p	Autor's research
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	*WorldBank*[1]
X44	Fiscal Space in t-p	Autor's calculations

Note: [1]A Cross-Country Database of Fiscal Space, 2019.
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 test for $p = 1, 2, 3$ for each feature. There are 44 variables included in 3 lagged so 132 features tested. Fiscal Space is measured as the difference between country public debt and EU median debt for each year.