Csaba Tóth G.

The Forecasting Capacity of Indicators Measuring Budget Sustainability

Summary: Since the outbreak of the European debt crisis, fiscal sustainability as a research area became more and more popular. This paper aims to assess the predicting power of the basic types of the following five fiscal sustainability indicators: primary gap (1), stationary tests for public debt (2), stationary test for the first differential of public debt (3), public revenues and expenditures cointegration (4); fiscal reaction function (5). The results indicate that predictive power is rather low for all indicators except for the first one. One reason for this may be that these indicators ignore the factors outside the range of fiscal policy. After comparing our result with previous research studies, we conclude that the results of these tests are heavily influenced by the choice of parameters, samples and the econometric methods used. For these reasons, it appears reasonable to evaluate and compare the predictive power of all fiscal sustainability indicators. By utilising the lessons learned, certain procedures should be standardised on the one hand, where this is possible, and on the other, methodology should be developed further in order to increase effectiveness.[1]

Keywords: fiscal policy, fiscal rules, sustainability indicators, primary gap, public debt
JEL codes: E60, E62, H60, H69

The financial crisis of 2007–2008 and the subsequent real economic crisis led to debt crises in several European countries within a few years' time. More than half a dozen countries were forced to face the fact that they were unable to satisfy their financial needs in the usual manner, namely from the market, and were forced to seek assistance from various international organisations. Though there was no actual sovereign default to speak of in the majority of these countries as creditors to date have been refunded in each case – with the exception of Greece –,[2] turning to and seeking out external assistance is, from a certain aspect, an admittance of the fact that the fiscal processes in place are no longer sustainable.[3] Though a wealth of literature deals with the definition and interpretation of sustainability, and the chapters on the sustainability of public finances are also numerous, hereinafter I will be following the approach by *Croce* and *Juan-Ramon* (2003), according to whom a given fiscal policy is sustainable only if it does not endanger a country's solvency in the future either. This latter is the government's capacity to meet its current debt repayment obligations without debt rescheduling or any other similar form of external assistance (Burnside, 2005). Based on the above, we can state that the assistance and contribution provided by international organisations to restore or sustain a country's solvency (through lending or any other form) is a sign or even a consequence of the fact that the given country's fiscal policy is unsustainable.

Such situations typically have grave consequences for countries: real economic

Email address: tothg@mnb.com

downturn, the introduction of forced austerity measures, increasing poverty, social conflicts and other costs (Reinhart – Rogoff, 2009; Vidovics-Dancs, 2014). At the same time, for science they also represent an opportunity to learn in order to increase the chances of preventing similar future events by coming to understand the social and economic processes leading up to the situation at hand. Through the examination of the data of a relatively large group of countries, they contribute to enabling an assessment of the efficiency and effectiveness of methods measuring fiscal sustainability.

In the present study, I will make an attempt to individually assess the forecasting capacity of the basic types of the following sustainability[4] examinations applied by international literature: the primary gap (1), the stationarity of the debt ratio (2), the stationarity of the first differential of the debt ratio (3), the cointegration of the revenue and the expenditure side (4) and the fiscal reaction function (5).

The driving force behind my examination is the fact that these methods are highly popular and are frequently part of analyses[5] that take an in-depth look at the sustainability of the budget of a given country, region or group of countries, as well as the fact that we have very limited knowledge pertaining to the effectiveness and forecasting capacity of these methods. For this very reason, my research aims to answer what would have happened if just before the onset of the crisis we were to have measured the sustainability of public finances using various sustainability examinations; and would the results thus arrived at have been in harmony with what actually happened after the crisis broke out. Did budget policy truly prove to be unsustainable in countries where the various examinations predicted or sustainable where results supported this? Beyond the presentation of results, the present study has two other objectives. Firstly, to highlight that discussing the forecasting capacity and the effectiveness of given methods and the debating of related findings is an important criterion of the wide-spread use of forecasts based on various methodologies. Secondly, to draw attention to the fact that the results of the forecasts I have examined and the accuracy of these forecasts depend greatly on the parameters selected during the examination.

It should be emphasised that due to the explosive expansion[6] of literature dealing with fiscal sustainability and the development of various *early warning systems*, the procedures I have examined form but a part of existing and widely-used methods. The foundation of one of the most important and most perspective research directions is the analysis framework developed by *Cottarelli* (2011) after the breakout of the debt crisis. In the case of the various countries, it classifies into three groups the risks accompanying debt refinancing: the level of main fiscal indicators and their long-term projection (1), (macro-economic, economic policy or other) shocks related to the baseline scenario that endanger the refinancing of public debt through the deterioration of fiscal outlooks (2), and other country-specific factors (such as external disequilibrium for instance) (3). It was based on this method that *Baldacci et al.* (2011b) defined two new complementary indicators which contribute to the deeper and more accurate understanding of fiscal sustainability. The fiscal vulnerability index indicates what the financial position of a given country is like compared to its earlier experiences and practice. This is complemented by the fiscal stress index, which measures the probability of refinancing risk developing into substantial fiscal crisis.

In addition to the appearance of new procedures, the methods examined here continue to play an important role in measuring fiscal

sustainability (Cottarelli – Escolano, 2014), and as such, I continue to consider the examination of their forecasting capacity to be topical and important. Within the various examinations, however, researchers use highly varied methodologies which in turn lead to markedly different results. Due to volume constraints, presenting and calculating all possibilities is not possible; therefore, I will only attempt to present the classic procedure that serves as the foundation of the various examinations. As the methodology has developed greatly in the recent period, one of the consequences of this solution is that the forecasting capacity of the examinations I have conducted does not necessarily represent the most effective forecasting within the given examination. However, continuing the research allows an opportunity in the future to reproduce and assess the forecasting capacity of methods using the most recent methodological developments.

As far as the rest of the paper is concerned, following a brief introduction to the methodology, in the interest of comparability I will in each case conduct the examination by taking the very last year before the European crisis (2007) as the last data point of the time series used. Then I will, on the one hand, compare the results with those of similar examinations found in relevant literature; and on the other, will attempt to quantify the forecasting capacity of the various methods using simple descriptive statistics. For this reason, I have split the range of examined countries (27 Member States excluding Croatia, which was last to accede to the European Union) into two groups. As they were forced to seek external assistance (typically IMF loans[7] or ECB bond purchases[8]), public finances proved to be unsustainable in Greece, Portugal, Ireland, Spain, Italy, Cyprus, Hungary, Latvia and Romania. I view the other countries as places where fiscal policy has proved to be sustainable.

PRIMARY GAP

More than a decade ago, *Manasse et al.* (2003) called attention to the fact that in itself, no single fiscal indicator is suitable to make forecasts concerning sovereign default, and this is presumably also true for projections dealing with sustainability. In spite of this, however, we can state for certain that the development of the debt ratio is closely linked to both topics through financing need. In order to present the concept of the primary gap, attributable to *Blanchard* (1990), we should take the following equation, which decomposes the change of the debt ratio in very simple fashion, as our starting point:

$$\Delta b_t = \frac{r_t - g_t}{1 + g_t} b_{t-1} - pb_t \qquad (1)$$

Where is the debt ratio (public debt/GDP), is the real interest rate, is the rate of real growth, and is the primary balance of the budget as a percentage of GDP. Let us assume that r and g are constant in time, therefore, we can discard the time indexes. The dynamic term in the focus of our current examination is the first term of the right side of the equation:

$$u = \frac{r - g}{1 + g} \qquad (2)$$

One of the main features of the relationship of the dynamic term, the primary balance and the debt ratio is that if we fix the first two, we arrive at the equilibrium point of public debt. As presented in detail by *Mellár* (2002) as well, depending on whether the primary balance is positive or negative and whether the dynamic term is positive or negative, we distinguish four different cases. If the dynamic term is positive, in other words, real interest rate exceeds the growth rate, the equilibrium point is negative for a deficit budget and positive for a surplus budget, but is stable in neither case. If the dynamic term is negative,

the equilibrium point is definitely stable with a value greater than zero for a deficit budget and lower for a surplus budget.

After having shown how the value of the dynamic term, its sign in particular, can impact the change in public debt, we can use the correlations we arrived at to examine sustainability. In order to do this, we should merge and solve the above two equations for zero:

$$ub_{t-1} = \qquad (3)$$

Equation (3) indicates that we can assign a so-called optimal primary balance to the given dynamic term and the initial debt ratio, which will stabilise the debt ratio. If the actual balance is not worse than the optimal balance, i.e. the gap is positive, public debt can be sustained; conversely, however, further intervention is required to sustain public debt.

As the next step, I will examine what we could have observed about the EU-27 before the crisis based on the primary gap. I calculated the real interest rate for period *t* by subtracting the deflator for period *t*+1 from the implicit interest rate of public finances for period *t*. It must be emphasised, however, that similarly to the nominal interest rate, the real interest rate in the equations used is also forward-looking, therefore, debt (change) in period *t* is (also) a function of real interest rate in period *t*–1.

The source of the data was in each case the database of the European Commission (AMECO). In the case of the various countries, I arrived at the dynamic term by averaging the value of dynamic terms between 1998 and 2007, a period the selection of which was on the one hand explained by the availability of ESA95 data, and on the other, that Blanchard himself also recommends 10-year averaging in his much quoted work (1990, p. 14).

According to my results, there were six countries where the actual value of the indicator in 2007 was worse than the (optimal) primary balance required to stabilise the debt ratio *(see Table 1)*. Of the six countries, Romanian, Greek, Hungarian and Portuguese budget policy did prove to be unsustainable in the end, while in contrast with forecasts, Great Britain and France required no external assistance. Of the countries deemed sustainable on the basis of the examination, Cyprus, Latvia, Italy, Ireland and Spain slipped into a financing crisis, while the forecasts concerning sustainability proved to be correct for all other countries.

In line with international literature, I will attempt to capture the efficiency of forecasting by means of three indicators. In the case of our examination, the sensitivity indicator (*true positive rate; TPR*)[9] shows how accurate the forecast was in countries which later proved to be unsustainable, in other words, it is primarily able to quantify Type II errors. As the primary gap indicated this beforehand for four of the nine countries that proved to be unsustainable, the value of the TPR is 45 per cent. This indicator is complemented by the false alarm rate *(false positive rate; FPR)*[10], which is linked to the Type I error, and the lower the indicator's value, the more accurate the forecast *(see Table 2)*. Since of the countries that were subsequently proved to be sustainable the primary gap only projected unsustainability (erroneously) in the case of two, the value of the FPR is 11 per cent.

The classification accuracy[11], which is arrived at as a ratio of correct forecasts to all forecasts and serves the comparison of the various examinations, is 74 per cent.

As a last step, I compared the results with calculations of earlier examinations conducted on the basis of similar methodology. The values I arrived at are in part identical to results found in the literature, and the differences can be traced back to the specific characteristics of the methodologies used. *Aristovnik* (2008) for instance, when examining North European

Table 1

EXAMINATION OF THE PRIMARY GAP (AS A PERCENTAGE OF GDP)

	pb (2007)	pb (OPT)	Difference		pb (2007)	pb (OPT)	Difference
Romania	–2.2	–0.6	–1.6	Slovenia	1.2	–0.1	1.3
Greece	–2.0	–0.5	–1.5	*Italy*	3.4	1.8	1.5
Great Britain	–0.6	0.3	–0.9	Netherlands	2.4	0.3	2.1
France	0.0	0.6	–0.7	*Ireland*	1.1	–1.3	2.4
Hungary	–1.0	–0.5	–0.5	Belgium	3.8	1.1	2.7
Portugal	–0.2	0.2	–0.4	Estonia	2.6	–0.3	2.8
Lithuania	–0.3	–0.4	0.1	Bulgaria	2.3	–1.0	3.3
Slovakia	–0.4	–0.6	0.2	Luxembourg	3.9	–0.2	4.2
Malta	0.9	0.7	0.2	*Spain*	3.5	–0.7	4.3
Poland	0.4	0.0	0.5	Sweden	5.3	0.1	5.3
Czech Republic	0.4	–0.2	0.6	Denmark	6.4	0.5	5.9
Latvia	0.0	–0.7	0.7	Finland	6.8	0.1	6.7
Austria	1.9	0.6	1.3	*Cyprus*	6.5	–1.0	7.5
Germany	3.0	1.7	1.3				

pb: primary balance as a percentage of GDP Countries where budget policy has proved to be unsustainable are marked in italics..

Source: own calculations based on AMECO data

Table 2

THE ACCURACY OF PRIMARY GAP FORECASTING

		Facts	
		Unsustainable	Sustainable
Forecast	Unsustainable	Correct classification (TP): 4	Type I error (FP): 2
	Sustainable	Type II error (FN): 5	Correct classification (TN): 16

Source: own calculations

countries, also deemed Greek and French fiscal policy to be unsustainable, while (in contrast with me) he relied on data from between 2003 and 2006 to calculate the indicator in question. Examining newly acceded countries, however, *Aristovnik – Berčič* (2007) came to different conclusions, which is probably due to the fact that they were working with earlier data (2001–2004). For the sake of completeness it should be mentioned that there have been research projects where the actual budget balance was not compared with the balance required to stabilise the debt ratio, but rather with an indicator needed to achieve a defined debt target (e.g. Maastricht 60 per cent) (Croce – Juan-Ramon, 2003; Hauner et al., 2007); however, this method differs from the classic examination of the budget gap to such an extent that comparison with its results serves no purpose here.

STATIONARITY AND COINTEGRATION

Though there are also certain voices of criticism in relation to this (Bohn, 2007), among the measurements concerning fiscal sustainability, those based on the stationarity of public debt and on the cointegration of the revenue and the expenditure side are still highly popular today (Miyazaki, 2014). Such examinations focus on the present-value borrowing constraint *(PVBC)*. In order to arrive at this, we must first construct the current budget constraint:

$$G_t+(1+r_t)B_{t-1}=R_t+B_t \qquad (4)$$

where G represents the state's primary expenditure (reduced by interest payment), R represents total revenue, B represents gross public debt, and r represents real interest rate. Shifting the equation by one period and solving it we arrive at the intertemporal budget constraint:

$$B_t=\sum_{s=1}^{\infty}\frac{R_{t+s}-G_{t+s}}{\prod_{j=1}^{s}(1+r_{t+j})}+\lim_{s\to\infty}\prod_{j=1}^{s}\frac{B_{t+s}}{(1+r_{t+j})} \qquad (5)$$

If the second expression on the right side of the equation is zero, this means that the present value of the current total debt is equal to the present value of future primary balances (surpluses). In the interest of easier empirical testing, there should be some other important transformations made. Let us assume that real interest rate is stationary, its average value is r, then define the following variable:

$$E_t=G_t+(r_t-r)B_{t-1} \qquad (6)$$

In this case, PVBC can be defined as follows:

$$B_{t-1}=\sum_{s=0}^{\infty}\frac{1}{(1+r)^{s+1}}(R_{t+s}-E_{t+s})+\lim_{s\to\infty}\frac{B_{t+s}}{(1+r)^{s+1}} \qquad (7)$$

It is a crucial requirement of sustainable fiscal policy that the second member of the right side of the equation should converge on zero on the infinite time horizon, as this is what prevents debt increasing at a rate greater than real interest rate. In other words, this is what ensures avoidance of the Ponzi scheme and compliance with the intertemporal budget constraint. Due to the transversality condition, the present value of the future primary balances of the government must be equal to the present value of public debt. In the case of growing economies, these correlations can also be expressed with variables as percentage of GDP, where y is the growth rate of gross domestic product:

$$\frac{B_t}{Y_t}=\frac{(1+r_t)}{(1+y_t)}\frac{B_t-1}{Y_t-1}+\frac{G_t}{Y_t}-\frac{R_t}{Y_t} \qquad (8)$$

Assuming that the real interest rate is stationary at an average r and that GDP growth is constant *(y)*, we can construct the following budget constraint

$$b_{t-1}=\sum_{s=0}^{\infty}\left(\frac{1+y}{(1+y)}\right)^{s+1}[p_{t+s}-e_{t+s}]+\lim_{s\to\infty}b_{t+s}\left(\frac{1+y}{1+r}\right)^{s+1} \qquad (9)$$

where $b_t=B_t/Y_t$; $p_t=R_t/Y_t$, $e_t=E_t/Y_t$. If $r>y$, that is real interest rate exceeds the GDP growth rate, the solvency condition is only satisfied if

$$\lim_{s\to\infty}b_{t+s}\left(\frac{1+y}{1+r}\right)^{s+1}=0$$

The result is similar to those before, in other words, fiscal policy can be deemed sustainable if the present value of future primary balances as a percentage of GDP is equal to the present value of the debt ratio. If the rate of real growth is greater than the real interest rate, then the process may be sustained even with a primary deficit of a certain size.

By using equation (7), we can measure the avoidance of the Ponzi scheme by checking whether the debt is stationary. To conduct

another examination, we should introduce another variable, namely total expenditure:

$$GG_t = G_t + r_t B_{t-1} \quad (10)$$

and define the budget constraint as follows:

$$GG_t - R_t = \sum_{s=0}^{\infty} \frac{1}{(1+r)^{s-1}} (\Delta R_{t+s} - \Delta E_{t+s}) + \lim_{s \to \infty} \frac{B_{t+s}}{(1+r)^{s+1}} \quad (11)$$

In order to avoid a Ponzi scheme, the variables *GG* and *R* must be cointegrated. To verify this, as a first step we must construct the following cointegration equation:

$$R_t = a + bGG_t + u_t \quad (12)$$

If the two variables are not stationary on their own, and the remainder term *(u)* featured in the regression equation contains no unit root, total revenue and total expenditure are cointegrated. In addition, the results also allow us to draw numerous other conclusions. In general, we can state that if there is no cointegration, fiscal policy is unsustainable (1). If there is cointegration, but $b<1$, fiscal policy is still unsustainable (2). If there is cointegration and $b=1$, fiscal policy is sustainable. (Afonso – Jalles, 2012).

At this point, it should be emphasised that important questions arise when completing the econometric examinations presented until this point. There are, for example, considerable methodological differences between the examinations related to the stationarity of debt that impact end results as well as the conclusions drawn. Some analyses focused on the first differential of public debt (Afonso – Rault, 2007; Afonso – Jalles, 2012; Talpos – Enache, 2008; Bohn, 1991), while another significant group on the level of public debt (Uctum – Wickens, 1996; Mahmood – Rauf, 2012; Artis – Marcellino, 2000; Uctum et al. 2006). As both solutions seem justifiable, I will examine both level and change.

Literature is also diverse in terms of the unit used to measure public debt. Earlier works frequently used real debt (Afonso – Rault, 2007), the debt-to-GNP ratio (Bohn, 1991), as well as the so-called discounted debt ratio (Uctum – Wickens, 1996), in recent years; however, a professional consensus seems to have been reached that the most appropriate indicator in terms of the examination is the debt-to-GDP ratio (Afonso – Jalles, 2011). This can be explained, on the one hand, by the fact that this is the indicator creditors primarily observe, and on the other, that it takes into account the country's income generating ability as well, which is of key importance with respect to sustainability. This is also what *Elmendorf – Mankiw* (1998) refer to when they call attention to the fact that an increase in debt does not necessarily indicate a Ponzi scheme. If real interest rate *(r)* is greater than the growth rate of the economy *(g)*, the process truly is unsustainable; however, conversely there is nothing to stop the government to finance interest rates from other loans as well. Precisely for this reason, I will also examine the changing of the debt ratio, and similarly, measure total revenue and total expenditure in relation to gross domestic product.

The third issue that we are unable to answer based on the literature is model fitting. As *Baumöhl et al.* (2011, p. 12) put it, "[n]evertheless, it is still questionable whether the trend stationarity is still a sufficient condition for fiscal sustainability." Given that the majority of examinations look at stationarity around the constant, and that based on the topic, and particularly due to the use of the debt-to-GDP ratio, this solution seems more justified, this is also what I will be examining in the case of the unit root tests.

THE STATIONARITY OF THE DEBT RATIO AND ITS FIRST DIFFERENTIAL

Similarly to the other econometric examinations conducted on fiscal indicators, the greatest problem is again the shortness of data series. In order to have the longest possible data series available for the examination of debt ratio stationarity, I have once again used the data of the European Commission (AMECO). Thus, I have managed to compile data series of at least 30 years for 13 Western European countries, and since the risk of drawing erroneous conclusions is significant with time series shorter than this, from this point I will only examine the data of the aforementioned 13 countries.

I tested the presence of stationarity using the augmented Dickey–Fuller test for both the level of the debt ratio and its first differential. The results arrived at based on the two examinations vary considerably.

When examining the debt ratio level, with 10 per cent significance, the presence of the unit root can be rejected for Belgium, Great Britain, Ireland and Portugal. In the case of the former two, the forecast concerning sustainability proved to be accurate; in the latter two, however, fiscal policy became unsustainable. Of the other 9 countries, forecasts concerning unsustainability were accurate in the case of Greece, Italy and Spain, but inaccurate for all the other countries *(see Table 3)*.

As the Type II error is relatively low, the sensitivity indicator (TPR) is 60 per cent, not far below the value of the previous method; however, due to the significant Type I error, the false alarm rate (FPR) is 75 per cent. Overall, the classification accuracy of forecasting based on debt ratio stationarity is 38 per cent *(see Table 4)*.

My results for the various countries are substantially different from the findings of *Claeys* (2007), which may be explained by the fact that his examinations used data from another source (OECD), covered a shorter

Table 3

THE VALUE OF TEST STATISTICS IN THE DEBT RATIO AND ITS FIRST DIFFERENTIAL (ADF-TEST)

	N	d	Δd
Austria	38	–2.19	–3.26**
Belgium	39	–3.28**	–1.99
Denmark	37	–1.71	–2.28
Finland	38	–1.24	–3.90***
France	31	–0.84	–3.48***
Greece	38	–0.34	–5.47***
Ireland	38	–3.11**	–1.46
Italy	38	–1.58	–2.83*
Netherlands	33	–2.25	–2.01
Portugal	35	–3.51***	–4.78***
Spain	38	–1.50	–1.98
Sweden	38	–1.81	–0.45
Great Britain	38	–3.4**	–3.95***

Note: *** –1% significance level; ** –5% significance level; * –10% significance level.

Source: own calculations based on AMECO data

period (1970–2001) and used a number of different tests (e.g. KPSS). The results I arrived at are less different from the findings of Baumöhl et al. (2011), though there were again a number of methodological differences, as they examined a different period (2000–2010) and worked with quarterly data.

Based on the examination of the first differential of the debt ratio, with 10 per cent significance, the presence of the unit root can be rejected for Austria, Finland, France, Great Britain, Greece, Italy and Portugal. This means that in the case of the first four countries, the forecast concerning sustainability proved to be accurate; the latter three, however, proved it wrong. Of the other six countries, fiscal policy did prove to be unsustainable in Spain and Ireland, while the forecast concerning unsustainability proved to be incorrect in the case of the other four countries.

The sensitivity indicator (TPR) reflecting the size of the Type II error is in this case 40 per cent, while the false alarm rate (FPR) is 50 per cent and accordingly, the classification accuracy for the projection based on the stationarity of the first differential of the debt ratio is 46 per cent *(see Table 5)*.

The results arrived at the level of the various countries are practically identical to those of *Prohl* and *Schneider* (2006), which is due to the fact that they worked from identical sources and according to identical methodology; however, their data series were a few years shorter (1970–2004). The latter cannot be said of the work of *Afonso* and *Jalles* (2011), whose data series covered close to 130 years (1880–2009). This is perhaps an explanation why our results are not identical, as they found that the first differential of the debt ratio of practically all the countries are stationary.

COINTEGRATION OF REVENUE AND EXPENDITURE

In recent years, of the examinations measuring fiscal sustainability, one of the most popular and most frequently used methods is based

Table 4

ACCURACY OF FORECASTING BASED ON THE STATIONARITY OF THE DEBT RATIO

		Facts	
		Unsustainable	Sustainable
Forecast	Unsustainable	Correct classification: 3	Type I error: 6
	Sustainable	Type II error: 2	Correct classification: 2

Source: own calculations

Table 5

ACCURACY OF FORECASTING BASED ON THE STATIONARITY OF DEBT RATIO CHANGE

		Facts	
		Unsustainable	Sustainable
Forecast	Unsustainable	Correct classification: 2	Type I error: 4
	Sustainable	Type II error: 3	Correct classification: 4

Source: own calculations

on the cointegration of the revenue and expenditure side. As before, I have once again used the data of the European Commission and have examined the same 13 countries. Total revenue and total expenditure are shown as a percentage of GDP in the database.

Since this examination requires I(1) processes in order to show cointegration indicating fiscal sustainability, as a first step I examined whether the data series separately contain unit roots *(see Table 6)*. The calculations indicate that in the case of Austria, Belgium, France and Great Britain, the null hypothesis cannot be clearly rejected, in other words, the time series of the expenditure and/or revenue side do not have unit roots, which leaves us with no reason to search for cointegration in the case of these countries. For the remaining nine countries, I examined whether the first differential of variables contained unit roots, and since the ADF test was unable to reject unit root presence for Greece, the Netherlands, Italy and Sweden, I was forced to narrow subsequent examinations to the remaining five countries. In their case, I examined cointegration using the Engle–Granger test, but as unit root presence could not be rejected in residues, I was forced to reject the possibility of cointegration for all countries.

Since we were unable to prove the presence of cointegration in the case of any of the countries, the results can only be managed if the forecast predicts fiscal unsustainability in each case *(see Table 7)*.

As a result, the sensitivity indicator (TPR) may be 100 per cent, but so is the false alarm

Table 6

THE RESULTS OF UNIT ROOT TESTS AND COINTEGRATION EXAMINATIONS

	Unit root tests				Engle–Granger	
	R	GG	ΔR	ΔGG	residue	b[x]
	Test -stat	Test -stat	Test -stat	Test -stat	Test -stat	coeff.
Austria	–2.32	–2.66*	–5.81***	–0.364	–2.83	0.70***
Belgium	–3.31**	–1.98	5.73***	–2.59*	–1.45	0.46***
Denmark	–1.51	–2.07	–4.98***	–3.90**	–2.14	0.67***
Finland	–1.99	–1.96	–4.85***	–3.30**	–2.12	0.69***
France	–2.68*	–2.51	–4.00***	–3.19**	–3.48**	0.75***
Greece	–0.60	–1.51	–1.98	–6.90***	–1.43	0.65***
Ireland	–2.09	–1.77	–6.57***	–4.19***	–1.80	0.23***
Italy	–1.28	–2.24	–5.62***	–1.89	–0.63	0.86***
Netherlands	–2.11	–1.60	–1.72	–5.68***	–4.71***	0.62***
Portugal	–1.37	–0.12	–4.79***	–4.65***	–1.98	0.74***
Spain	–1.90	–1.76	–2.76*	–3.37***	–0.19	0.80***
Sweden	–1.99	2.05	–1.17	–3.54**	–3.56**	0.48***
Great Britain	–3.25**	–3.45***	–2.03	–1.99	–3.63**	0.29***

x In cointegration equation (12), this is the total expenditure parameter, where the dependent variable was total revenue. *** –1% significance level; ** –5% significance level; * –10% significance level.

Source: own calculations based on AMECO data

rate (FPR), and the classification accuracy that best captures the effectiveness of cointegration-based forecasting is 38 per cent.

The results should be compared with those of the examination of Afonso (2005), conducted on a practically identical group of countries. He also judged the majority of old EU countries unsustainable, but in the case of five Member States (Austria, the Netherlands, Germany, Finland, Portugal) he showed cointegration between the revenue and the expenditure side. A possible explanation for the different (and in terms of forecast accuracy, better) result is the allowance of the structural break, but in addition to the ADF test, the use of the Phillips–Perron (PP) test may also have contributed. The same seems to be true for both the differences between results, and the possible explanations if I compare my results with newer examinations by Afonso and Jalles (2012).

FISCAL REACTION FUNCTION

One of the most widely-used type of sustainability examinations is attributed to *Bohn* (1998), who was the first to analyse US budget data using the reaction function-based method. The essence of the procedure is the examination of the correlation between two (or more) variables. One must be a fiscal instrument that indicates the changes of economic policy, while the other must reflect fiscal goals. In the case of research examining the sustainability of public debt, the obvious objective is to maintain the stability of public debt, while the other (fiscal) variable in the correlation is the primary balance. Many have studied the impact of fiscal policy on public debt, be it an analysis of debt dynamics (Hall – Sargent, 2010)[12], or even the impact of budget policy on interest rates (Ardagna – Caselli – Lane, 2004). Bohn, however, drew attention to the fact that primary balance is not the only thing that can impact public debt, the effect may be mutual and is actually closely related to sustainability. If, through the primary balance, a government responds swiftly and efficiently to the changing of public debt, it practically nips the danger of public debt becoming unsustainable in the bud. Accordingly, in the case of examinations based on the reaction function, we deem public debt (and underlying economic policy) sustainable if it can be proved with historical data that in response to the increase of public debt, the position of the budget improves, and it prevents (or at least slows) indebtedness, whereas it is deemed unsustainable if budget policy is inflexible to the changing of the debt ratio.

In order to measure the fiscal reaction function, the estimation of the following regression equation is required:

$$pb_t = const + \alpha b_{t-1} + \beta pb_{t-1} \qquad (13)$$

where pb is the cyclically adjusted primary budget balance measured as a proportion of

Table 7

THE ACCURACY OF COINTEGRATION-BASED FORECASTING

		Facts	
		Unsustainable	**Sustainable**
Forecast	Unsustainable	Correct classification: 5	Type I error: 8
	Sustainable	Type II error: 0	Correct classification: 0

Source: own calculations

potential GDP, and b is the debt-to-GDP ratio. The objective of cyclical adjustment is to filter out non-discretionary measures. If we want to know how decision-makers react to the changing of debt, we must disregard the effect of balance changes due to cyclical changes (such as the automatic rising of tax revenues for example).[13] A part of the relevant literature resolves this by featuring the output gap among explanatory variables (Izak, 2009). The use of a cyclically adjusted balance indicator seems a more accurate solution (IMF, 2004; Gali – Perotti, 2003).

As before, the source of the debt ratio used in the examination is the database of the European Commission (AMECO), while in the case of cyclically adjusted primary balances this source is the OECD (1996, 2012), as this latter has longer data series of this particular indicator. Through the two data sources, I have an at least 30 year database at my disposal with respect to the 13 countries examined earlier, therefore, I will examine sustainability for these countries until 2007. As a first step, using the regression equation I will prepare an estimate of the debt ratio parameter (α), then check whether the residue behaves as white noise or is normally distributed.

As revealed by Table 4, which shows the results, the debt ratio parameter is significant for Austria, Belgium, Ireland, Italy, Spain, Portugal, Denmark and Sweden alike. In the case of the latter three countries, however, we were forced to reject the relevant results of our model calculation either due to the normality test or the Ljung–Box test. All this means that the method was accurate in predicting sustainability for Austria and Belgium, but inaccurate in the case of Ireland, Italy and Spain. It qualified the other eight countries as unsustainable, a prediction that proved to be true in the case of Portugal and Greece, but false for the other six countries *(see Table 8).*

The sensitivity indicator (TPR), which can be traced back to the Type II error, is 40 per cent; the false alarm rate (FPR), related to the numerous Type I errors is 75 per cent; while the classification accuracy is 31 per cent *(see Table 9).*

My results concerning the various countries examined are only in part identical to the calculations of the IMF (2004), which covered nearly the same period (1971–2003), as those calculations, with the exception of Ireland, Germany and France, deemed all states sustainable. The difference may be due to a number of factors. Firstly, monetary fund experts failed to examine the distribution of residues; secondly, they featured the difference of the short-term interest rate and the benchmark rate calculated on the basis of the rule describing monetary policy among the explanatory variables of the regression equation, which difference is meant to quantify the so-called 'monetary gap'; and thirdly, they took the possibility of structural break into account in the regression in the case a number of countries.

CONCLUSIONS

Of the examinations presented in this study, somewhat surprisingly the forecasting capacity of the method that is viewed as considerably more static than the others (primary gap) proved to be the most accurate. As part of this, I compared the primary balance required to stabilise debt with the current (2007) balance, in other words, instead of a longer time series, I had to determine sustainability on the basis of a selected year. In my view, the relatively (!) good forecasting capacity can be attributed to the fact that during the time of the financial crisis, current fiscal equilibrium plays an increased role in investor opinions, and the balance of the last year before the crisis plays

Table 8

PARAMETERS OF THE FISCAL REACTION FUNCTION

	α	Residual chi-square	Ljung-Box Q
Austria	0.028**	1.207	1.633
Belgium	0.041***	0.076	0.721
Denmark	0.029***	2.722	3.094*
Finland	0.019	3.766	0.047
France	0.002	0.86	1.401
Greece	0.009	0.987	0.107
Ireland	0.033***	1.18	0.287
Italy	0.098***	0.622	0.467
Netherlands	0.031	8.796**	0.027
Portugal	0.058*	12.903***	0.002
Spain	0.015*	1.68	0.373
Sweden	0.042*	12.357***	1.469
Great Britain	0.006	0.19	7.102***

Note: *** –1% significance level; ** –5% significance level; * –10% significance level.

Source: own calculations based on AMECO data

Table 9

ACCURACY OF FORECASTING BASED ON THE FISCAL REACTION FUNCTION

		Facts	
		Unsustainable	Sustainable
Forecast	Unsustainable	Correct classification: 2	Type I error: 6
	Sustainable	Type II error: 3	Correct classification: 2

Source: own calculations

an important role in this examination *(see Table 10)*.

At the same time, the forecasting capacity of the other examinations and the classification accuracy failed to reach 50 per cent, which can be attributed to a number of reasons. On the one hand, the structural processes and problems of the budget cannot necessarily be captured using the four indicators (revenue, expenditure, balance, debt) in focus. On the other, fiscal sustainability is impacted by numerous factors that fall outside of public finances (Manasse – Roubini, 2005). In the case of Spain and Ireland, for instance, the financing problem was not caused directly by budgetary processes, but rather by the necessary bailouts to certain commercial banks in trouble (Stein, 2011). The premise behind the methods used also certainly contributed to the low forecasting capacity of examinations. This treated the performance of fiscal governance during the entire period under review in a uniform manner, and this is also what it projected to the future. The

Table 10

THE EFFECTIVENESS OF THE VARIOUS FORECASTING METHODS

Examination	Classification accuracy	Sensitivity (TPR)	False alarm (FPR)
Primary gap	74%	44%	11%
Stationarity of the debt ratio change	46%	40%	50%
Stationarity of the debt ratio	38%	60%	75%
Cointegration of the revenue and the expenditure side	38%	100%	100%
Fiscal reaction function	31%	40%	75%

Source: own calculations

process was also reinforced by the fact that in addition to government behaviour, we also assumed external conditions to be constant, and disregarded the very important lesson that a fiscal policy that is sustainable in 'peacetime', may prove to be unsustainable amidst a financing and confidence crisis.

A comparison of results with literature suggests that forecasts and forecast accuracy change substantially depending on the answers given to fundamental methodological questions. This is true for the measurement units selected, the frequency of data series, the indicators, the estimation methods, and the procedures used to test hypotheses. Furthermore, it is especially true for the start and length of the periods reviewed. In addition, within a given examination, the procedures selected were typically based on the simplest methods, which were least developed in terms of methodology. The examination of structural breaks, for instance, would have been beyond the scope exceeded the limits of this research, and though it would have improved the accuracy of certain forecasts (cointegration, fiscal reaction function) (Afonso – Rault, 2008), it would have required longer time series, which is particularly problematic in the field of budgetary indicators.

All in all, the results achieved are in line with the findings of studies that were conducted using different methods and indicators, but which also focused on the forecasting capacity of various indicators (Berti et al., 2012; De Cos et al., 2014). These few studies conclude that the inclusion of country-specific factors (threshold values) in the examination increases efficiency, that it is expedient to supplement the range of fiscal indicators with non-fiscal indicators, and that the forecasting capacity of complex indexes generated using various methods is better than the predictive power of certain indicators.

This confirms the need to use discretion with the results of forecasts based on the indicators I have examined. This does not mean that these examinations are not suitable for forecasting or that there is no need for such methods to measure fiscal sustainability, quite the opposite in fact. However, to ensure that decision-makers are provided with results obtained from examinations of truly the best forecasting capacity, I consider it important to make a broader comparison of the forecasting capacity of various methods (Baldacci et al., 2011a). By utilising the lessons learned, certain procedures should be unified on the one hand, where this is possible, and on the other, methodology should be developed further in order to increase effectiveness.

Notes

[1] The author would hereby like to thank Gergely Baksay, András Balatoni, Gábor P. Kiss and Zsanett Orlovits for their useful advice, and would also like to say thank you to the speakers at the conference organised by the Hungarian Society of Economists in December 2013, Ágnes Vidovics-Dancs in particular. The author shall of course claim all responsibility for any errors and mistakes that may occur in the study.

[2] For more on the definitions related to sovereign default, see Vidovics-Dancs (2013)

[3] t must also be emphasised of course that the sustainability of the budget is determined by future budget policy and as such sustainability, in the literal sense of the word, cannot be measured (Pápa – Valentinyi, 2008)

[4] In the following, by sustainability I will be referring to the sustainability of the budget.

[5] For example, IMF (2004), Afonso – Rault (2007), Prohl – Schneider (2006), Claeys (2007), Aristovnik (2008), Baumöhl et al. (2011), Talpos – Enache (2008), Miyazaki, (2014).

[6] see, for instance, IMF (2010)

[7] http://www.imf.org/external/np/exr/facts/europe.htm

[8] http://tozsdeakademia.com/piaci-hirek/az-ekb-spanyol-es-olasz-kotvenyeket-vasarol

[9] TPR=TP/(TP+FN)

[10] FPR=FP/(FP+TN)

[11] classification accuracy=(TP+TN)/(TP+TN+FP+FN)

[12] On Hungarian data, see Tóth (2012)

[13] For more on the role and diversity of fiscal indicators, see P. Kiss (2011), and for more on cyclical adjustment, see Hoffmann – P. Kiss (2010).

Literature

Afonso, A. (2005): Fiscal Sustainability: The Unpleasant European Case. FinanzArchiv, 61 (1), pp. 19–44 Afonso, A. – Jalles, J. T. (2011): A Longer-Run Perspective on Fiscal Sustainability. Working Papers 2011/17, Department of Economics at the School of Economics and Management (ISEG), Technical University of Lisbon

Afonso, A. – Jalles, J. T. (2012): Revisiting Fiscal Sustainability: Panel Cointegration and Structural Breaks in OECD Countries. *Working Paper Series 1465, European Central Bank*

Afonso, A. – Rault, C. (2007): What Do We Really Know about Fiscal Sustainability in the EU? A Panel Data Diagnostic. *ECB Working Paper* 820

Afonso, A. – Rault, C. (2008): Should we Care for Structural Breaks when Assessing Fiscal Sustainability? *Working Papers Department of Economics* 2008/01, ISEG - School of Economics and Management, Department of Economics, University of Lisbon.

Ardagna, S. – Caselli, F. – Lane, T. (2004): Fiscal Discipline and the Cost of Public Debt Service: Some Estimates for OECD Countries. *ECB Working Paper* 411. Frankfurt am Main

Aristovnik, A. – Berčič, B. (2007): Fiscal Sustainability in Selected Transition Countries. Ekonomický časopis. Bratislava, *Slovak Akad. Press*, 7.55: pp. 659–675

Aristovnik, A. (2008): Fiscal Sustainability in the Mediterranean Region – A Comparison Between the EU and Non-EU Member States. *Romanian Journal of Economic Forecasting* Vol. 9, Issue 4, pp. 161–173

Artis, M. – Marcellino, M. (2000): The Solvency of Government Finances in Europe. *Banca d'Italia* pp. 209–241

Baldacci, E. – Iva, P. K. – Nazim, B. – Gabriela, D. – Samah, M. (2011a): Assessing Fiscal Stress. *IMF Working Papers*. WP/11/100

Baldacci, E. – McHugh, J. – Petrova, I. (2011b): Measuring Fiscal Vulnerability and Fiscal Stress: A Proposed Set of Indicators. *IMF Working Paper* 11/94.

Baumöhl, E. – Výrost, T. – Lyócsa, Š. (2011): Are we Able to Capture the EU Debt Crisis? Evidence from PIIGGS Countries in Panel Unit Root Framework. *MPRA Paper 30334, University Library of Munich, Germany*

Berti, K. - Salto, M. - Lequien, M. (2012): An Early-Detection Index of Fiscal Stress for EU Countries (No. 475). *Directorate General Economic and Monetary Affairs (DG ECFIN), European Commission*

Blanchard, O. (1990): Suggestions for a New Set of Fiscal Indicators. *Economics Department Working Paper 79. Paris: OECD*

Bohn, H. (1991): Budget Balance Through Revenue or Spending Adjustments? Some Historical Evidence for the United States. *Weiss Center Working Papers* 3–91, Wharton School – Weiss Center for International Financial Research

Bohn, H. (1998): The Behavior of U.S. Public Debt and Deficits. *The Quarterly Journal of Economics.* Vol. 113. No. 3. pp. 949–963

Bohn, H. (2007): Are Stationarity and Cointegration Restrictions Really Necessary for the Intertemporal Budget Constraint? *Journal of Monetary Economics.* Elsevier, vol. 54(7), pp. 1837–1847

Burnside, C. (ed.) (2005): Fiscal Sustainability in Theory and Practice: A Handbook. *The WorldBank Publications.* Washington D.C.

Claeys, P. (2007): Sustainability of EU Fiscal Policies: A Panel Test. *Institut de Recerca en Economia Aplicada, University of Barcelona, Documents de Treball, 02.*

Cottarelli, C. (2011): The Risk Octagon: a Comprehensive Framework for Assessing Sovereign Risks. *Presentation at the Sapienza University in Rome on January 25*

Cottarelli, C., – Escolano, J. (2014): Debt Dynamics and Fiscal Sustainability. *Post-Crisis Fiscal Policy, pp. 31–47 The MIT Press, Cambridge*

Croce, E. – Juan-Ramon, H. V. (2003): Assessing Fiscal Sustainability: A Cross-Country Comparison. *International Monetary Fund Working Paper.* No. 145. Washington D.C.

De Cos, P. H. – Moral-Benito, E. – Koester, G. B. – Nickel, C. (2014): Signalling Fiscal Stress in the Euro Area: A Country-Specific Early Warning System. *ECB Working Paper* 1712.

Elmendorf, D. W – Mankiw, N. G. (1998): Government Debt. In Handbook of Macroeconomics 1, part 3. pp. 1615–1669

Gali, J. – Perotti, R. (2003): Fiscal Policy and Monetary Integration in Europe. *Economic Policy*. Vol. 37 (October), pp. 533–572

Hall, G. J. – Sargent, T. J. (2010): Interest Rate Risk and Other Determinants of Post-WWII U.S. Government Debt/GDP Dynamics. *National Bureau*

of Economic Research Working Paper Series. No. 15702. Cambridge

Hauner, D. – Leigh, D. – Skaarup, M. (2007): Ensuring Fiscal Sustainability in G7 Countries. *IMF Working Papers, 07/187*

Hoffmann, M. – P. Kiss, G. (2010): From those Lying Facts to the Underlying Deficit, *MNB Bulletin, National Bank of Hungary*, vol. 5(4), pp. 7–16

Mahmood, T. – Rauf, S. A. (2012): Public Debt Sustainability. Evidence from Developing Country. *Pakistan Economic and Social Review* 50.1 pp. 23–40

Manasse, P. – Roubini, N – Schimmelpfennig, A. (2003): Predicting Sovereign Debt Crises. *IMF Working Paper* 03/221

Manasse, P. – Roubini, N. (2005): „Rules of Thumb" for Sovereign Debt Crises. *IMF Working. Papers, N. 05/42, March*

Mellár, T. (2002): Néhány megjegyzés az adósságdinamikához (A Few Notes on Debt Dynamics). *Közgazdasági Szemle (Economic Review)*. Vol. XLIX, pp. 725–740

Miyazaki, T. (2014): Fiscal Reform and Fiscal Sustainability: Evidence from Australia and Sweden. *Discussion Papers 1407, Graduate School of Economics, Kobe University*

Pápa, L. – Valentinyi, Á. (2008): Költségvetési fenntarthatóság (Budget Sustainability). *Közgazdasági Szemle (Economic Review)*. Vol. LV, Issue 5, pp. 395–426

Prohl, S. – Schneider, F. G. (2006): Sustainability of Public Debt and Budget Deficit: Panel Cointegration Analysis for the European Union Member Countries. No. 0610. *Working Paper, Department of Economics, Johannes Kepler University of Linz*

P. Kiss, G. (2011): Mozgó célpont? Fiskális mutatók jegybanki szemszögből. (Moving Target? Fiscal Indicators from a Central Bank Perspective.) *MNB Studies 92.*

Reinhart, C. M – Rogoff, K. S. (2009): This Time is Different. *Princeton University Press. Princeton and Oxford*

Stein, J. L. (2011): The Diversity of Debt Crises in Europe. *CESIFO Working Paper.* No. 3348. Munich

Talpos, I. – Enache, C. (2008): Fiscal Policy Sustainability in Romania. *Annales Universitatis Apulensis Series Oeconomica*, Faculty of Sciences, "1 Decembrie 1918" University, Alba Iulia, vol. 1(10), pp. 233–243

Tóth, G. Cs. (2012): Debt Dynamics and Sustainability. *Hungarian Statistical Review*. Special Number 16., pp. 123–149

Uctum, M. – Wickens, M. (1996): Debt and Deficit Ceilings, and Sustainability of Fiscal Policies: an Intertemporal Analysis. *Federal Reserve of New York, Research Paper* No. 9615

Uctum, M. – Thurston, T. – Uctum, R. (2006): Public Debt, the Unit Root Hypothesis and Structural Breaks: A Multi-Country Analysis. *Economica* 73, pp. 129–156

Vidovics-Dancs, Á. (2013): Államcsődök. Tények és alapfogalmak újragondolva (Sovereign defaults. Facts and Basic Concepts Rethought). *Hitelintézeti szemle (Credit Institution Review). Vol.* 12, Issue 4, pp. 285–305

Vidovics-Dancs, Á. (2014): Államcsődök az államcsőd költségei régen és ma (Sovereign defaults – The Costs of Sovereign Default in the Past and Today). *Közgazdasági Szemle (Economic Review). Vol.* 61, Issue 3, pp. 262–278

IMF (2004): Has Fiscal Behavior Changed Under the European Economic and Monetary Union? Chapter II in World Economic Outlook, September 2004: The Global Demographic Transition, World Economic and Financial Surveys (Washington: International Monetary Fund).

IMF (2010): The IMF–FSB Early Warning Exercise: Design and Methodological Toolkit. International Monetary Fund, September 2009, (Washington DC). Available via the Internet: www.imf.org/external/np/pp/eng/2010/090110.pdf

OECD (1996): OECD Economic Outlook No. 60. OECD Economic Outlook: Statistics and Projections (database). Downloaded on: 09.08.2013

OECD (2012): OECD Economic Outlook No. 91. OECD Economic Outlook: Statistics and Projections (database). Downloaded on: 09.08.2013