



No. 04-2021
Jan. 2021

Fiscal Sustainability Hypothesis Test in Central and Eastern Europe: A Panel Data Perspective

Benjamin Owusu*

**Bielefeld University
P.O. Box 10 01 31
33501 Bielefeld – Germany**

ISSN 2196-2723

→ www.wiwi.uni-bielefeld.de

Fiscal Sustainability Hypothesis Test in Central and Eastern Europe: A Panel Data Perspective

Benjamin Owusu*

Abstract

This paper assesses the fiscal sustainability hypothesis for 10 Eastern and Central European countries (CEEC) between 1997 and 2019. The study adopts very recent panel econometric techniques which accounts for issues of structural breaks and cross-sectional dependence in the data generating process to examine the cointegration between government revenue and expenditures. Preliminary results show that revenues and expenditures do not have a long-run relationship and hence a rejection of the sustainability hypothesis. As a next step, we discriminate between structural and cyclical components of revenues and expenditures in order to place emphasis on the structural component. We argue that the structural component of fiscal variables represents the actual long term behaviour of the policymaker. Further results indicate that structural revenues and expenditures have a long-run relationship however with a slope coefficient less than unity which implies sustainability in the weaker sense. At that point, expenditures exceed revenues and if this continues for a long time the government may find it difficult to market its debts in the long run. This result suggests that the fiscal authorities in CEEC must therefore do more by taking long term actions to counteract the rising fiscal deficit problems.

JEL: H0, H6, E6

Keywords: Fiscal Sustainability, Cointegration, Government Revenue, Government Expenditure

*Department of Business Administration and Economics, Bielefeld University, P.O. Box 100131, 33501 Bielefeld, Germany, e-mail: benjamin.owusu@uni-bielefeld.de

The author would like to thank Alfred Greiner and Dietmar Bauer for valuable comments that helped to improve the paper. Insightful comments by Tobias Becker and Henrietta Acquah-Swanzy is also acknowledged.

1 Introduction

The recent financial crises and global economic downturn prompted governments interventions across the world by way of fiscal expansions in attempts to stimulate aggregate demand ¹. This has implications on fiscal policy since spending has to be financed by public deficits (Greiner and Fincke, 2015). Rising public deficits and debts to unsustainable levels may have long-run implications for the government since holders of government debts (usually the private sector) could lose confidence in government bonds. Secondly, the government could also default on its debts if it reach unsustainable levels. The need to finance public deficit imposes a constraint on fiscal policy since governments in dynamically efficient economies have borrowing limits and also face a present value borrowing constraint.² The issue of fiscal sustainability has therefore received considerable attention both in theoretical and empirical discussions. The fiscal stance is said to be sustainable if the future total discounted primary surplus in present value terms is equal to current debt. In other words, future stream of primary surplus when discounted in present value terms should be sufficient enough to offset the current level of public debt. Violating the conditions of the Intertemporal Budget Constraint (IBC) implies that debt will soar to an unsustainable level at a faster rate than the growth rate of the economy.

Prior to the accession of the European Union (EU), governments in Central and Eastern European Countries (CEEC) had to institute extensive fiscal policy actions to adjust their budgets and transform structures of revenues and expenditures whilst implementing institutional frameworks for fiscal policy reforms (Gleich, 2003). The objective was to ensure that they meet the necessary fiscal criterion in terms of size of debts, deficits and other obligations as stipulated in the Maastricht Treaty (MT) and Stability and Growth Pact (SGP). Eight CEEC out of the ten countries that joined the EU from the so-called eastern enlargement scheme had lower debt to GDP ratios below the 60% threshold required by the MT and SGP, hence Hallet and Lewis (2007) speculated that these CEEC could follow an explosive debt path for years without necessarily violating the fiscal sustainability requirements. Sixteen years after joining the EU, it remains to be seen if indeed these countries have pursued sustainable fiscal policies.

Most pioneered literature on fiscal sustainability started by empirically testing the stationarity of government debt and deficits (Westerlund, 2010) as a way of fulfilling the government budget constraint. Notable among them are Hamilton and Flavin (1986), Trehan and Walsh (1988), Kremers (1988), Wilcox (1989) and Baglioni and Cherubini (1993). Later authors such as Hakkio and Rush (1991), Lui and Tanner (1995), Quintos (1995), Ahmed and Rogers (1995) and more recently Afonso (2005) and Westerlund and

¹This principle is based on the Keynesian concept which has dominated the political-economic principles lately. The Keynesian school of thoughts advocates for government intervention in order to stabilize market economies.

²See Abel et al. (1988) for a detailed discussion regarding dynamic efficiency of an economy

Prohl(2010), have all zoomed in to specifically consider a more flexible approach of the cointegration between government revenue and expenditures. This is to ascertain if indeed revenues and expenditures have a long-run relationship with a positive cointegration vector, which is a confirmation of the sustainability hypothesis.

Even though a vast stream of empirical studies on fiscal sustainability on the European continent have been undertaken, however there exist limited studies in the context of CEEC (Bökemeier and Stoian, 2018). For instance, Krajewski et al. (2016) examined the public debt sustainability for 10 selected CEEC countries using panel stationarity, a cointegration technique and a fiscal reaction function for the period 1990-2012. Their results indicated that the fiscal stance of selected CEEC countries is jointly sustainable. Similarly, Llorca and Redzepagic (2007) assessed the sustainability of fiscal policy for 8 CEEC countries using panel cointegration analysis and found out these countries pursued sustainable fiscal policies for period 1999 to 2006 using quarterly data. Bökemeier and Stoian (2018) also investigated debt sustainability in 10 CEEC countries using estimates of a fiscal reaction function in its cubic form over the period 1998 - 2015. Their results revealed that government debts were at sustainable levels and that governments had not reached fiscal fatigue thresholds. Even though the studies above employed panel sustainability test for CEEC, none incorporated the possibility of structural breaks and cross sectional dependence in the panel data generating process³.

The aim of this paper is to ascertain the fiscal sustainability of 10 CEEC for the period 1997 to 2019 by investigating the long-run relationship between revenues and expenditure using panel cointegration.⁴ The study makes use of a panel data analysis in order to benefit from the rich dynamism of panels. The availability of large macroeconomic datasets, over a long period of time and for different economies is a recipe for a shift in the mean or trend of the individual time series. This increases the probability of break occurrence in the data (Carrion-i-Silvestre et al, 2005). In cointegration analysis, structural changes have the tendency of affecting the cointegration vector, which is in contrast to conventional wisdom considering the fact that cointegration is a long-run stable relationship (Westerlund and Edgerton, 2008). This also leads to wrong inferences and hence proves how important it is to account for structural changes in the data generating process (Bai and Perron, 1998 ; Carrion-i-Silvestre et al, 2005).

Furthermore, cross country macroeconomic and financial datasets are associated with cross-sectional dependence because of inter-country links and dependencies, which is more of the rule now than an exception (Westerlund and Edgerton, 2008). Cross-sectional dependence affects the size properties of the unit root test and hence render inferences

³The accession of some countries to the EU or Eurozone could represent a structural change in policy due to requirements that must be met and maintained by members of the union, notably requirements enshrined in the so called MT and SGP.

⁴These countries are Czechia, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia, Bulgaria and Romania.

incredible.⁵ Hence this study adopts the so-called "second generational" econometric procedure which accounts for both cross-sectional dependence and structural breaks simultaneously in the data generating process unlike other previous panel studies such as Beqiraj et al (2018), Claeys (2007) and Llorca and Redzepagic (2007).

Preliminary results show that revenues and expenditures do not have a long term relationship and hence a rejection of the sustainability hypothesis. Further, we discriminate between structural and cyclical components of revenues and expenditures in order to place emphasis on the structural component. This is the novelty of this paper unlike previous panel cointegration sustainability studies between revenues and expenditures such as Westerlund (2010), Afonso (2005), Quintos (1995), Prohl and Schneider (2006), Claeys (2007), Llorca and Redzepagic (2007). With motivations from Gali et al. (2003), who posited that the component of fiscal variables whose variations do not emanate from the influence of cycles represents discretionary fiscal policy. As opined by Blanchard (2006), structural fiscal variables provide a benchmark by which fiscal policy can be judged. We argue that this structural component of fiscal variables (cyclically adjusted variables) represents the actual long term behaviour of the policymaker and should be examined when conducting sustainability analysis. Further results indicate that the cyclically adjusted revenue and expenditure have a long run relationship. However the slope coefficient of the cointegration relationship is less than unity and not strong enough to infer sustainability in the strong sense for cyclically adjusted variables⁶. These results suggest that even though cointegration exists for cyclically adjusted variables, the magnitude of the cointegration slope implies that expenditures are rising faster than revenue which indicates fiscal deficits. Hence debt to GDP ratio is not bounded and if this continues for a long time the debt stock will no longer be finite or sustainable implying a weaker form of sustainability. There is therefore the need for the fiscal authorities in the selected countries to pursue long term actions that counteract rising fiscal deficits by way of fiscal consolidation in order to ensure the satisfaction of the government IBC. The contribution of the paper is in three folds. Firstly, it employs recent advances in panel econometrics that models structural breaks and cross-sectional dependence simultaneously in the data generating process for the sustainability hypothesis test in CEEC. Secondly, the study makes a case for the use of structural fiscal variables which is devoid of automatic response variables in the cointegration analysis for the sustainability hypothesis test. Finally, the study adds to the growing literature on fiscal sustainability in CEEC region.

The rest of the study is structured as follows. Chapter 2 will discuss the methodology used for the paper by laying the theoretical foundations for the sustainability test. The chapter will further discuss the various econometric procedures used for the test. Chapter

⁵Banerjee et al (2004) argued that unit root test which assumes cross-sectional independence suffers from size distortions as the actual size of the test is lower than the empirical size.

⁶To infer strong sustainability in the sense of Quintos (1995), the cointegration slope must be equal to or greater than unity.

three will provide the empirical estimation and discussion of the results. Chapter four will finally conclude the paper.

2 Methodology

We begin with the government budget constraint which is assumed to hold at all times. A period government budget constraint in nominal terms is written as

$$g_t + i_t b_{t-1} = r_t + b_t - b_{t-1} \quad (1)$$

Where g_t represents government spending, b_t is the government bond, i_t is the interest rates on bonds and r_t represents the government revenue. The above equation show that the government expenditure (LHS) must be equal to total government receipts (RHS) at all times in order for the budget constraint to hold intertemporarily. Here we rule out the possibility of monetizing of spending or the activities of monetary authorities. That is, we do not consider government printing money (also known as seignorage) to fund its expenditure as this is known to cause inflation. This assumption is plausible because the characteristic of modern economies is such that central banks independently control monetary policy (Greiner and Fincke, 2015) with little or no influence from fiscal authorities.

Taking the state of the economy into consideration and assuming i_t to be stationary around its mean i , equation (1) can be re-written as

$$\frac{g_t}{y_t} + \frac{(1+i)b_{t-1}}{y_t} = \frac{r_t}{y_t} + \frac{b_t}{y_t} \quad (2)$$

where y_t represents national income or nominal GDP. Simplifying further leads to

$$\frac{g_t}{y_t} + \frac{(1+i)b_{t-1}}{y_t} \cdot \frac{y_{t-1}}{y_{t-1}} = \frac{r_t}{y_t} + \frac{b_t}{y_t} \quad (3)$$

$$\frac{g_t}{y_t} + \frac{(1+i)}{(1+f)} \cdot \frac{b_{t-1}}{y_{t-1}} = \frac{r_t}{y_t} + \frac{b_t}{y_t} \quad (4)$$

where f is the nominal growth rate of the economy (GDP). Using capital notations, (4) can be rewritten as

$$G_t + (1+\rho)B_{t-1} = R_t + B_t \quad (5)$$

Where $\frac{g_t}{y_t} = G_t$, $\frac{b_{t-1}}{y_{t-1}} = B_{t-1}$, $\frac{r_t}{y_t} = R_t$, $\frac{b_t}{y_t} = B_t$. $\rho = \frac{i-f}{1+f}$ is the growth adjusted interest rate which is assumed to be stationary for sake of simplicity.

Further modification required for empirical estimation, let $G_t^r = G_t + (\rho_t - \rho)B_{t-1}$, where ρ is the mean real interest rate and stationary.

Assuming that (5) holds continuously, then by forward substitution (see appendix 5.1 for proof), the present value budget constraint can be written as

$$B_{t-1} = \sum_{s=t+1}^{\infty} \left(\frac{1}{1+\rho} \right)^{s-t} (R_{s-1} - G_{s-1}^r) + \lim_{s \rightarrow \infty} \left(\frac{1}{1+\rho} \right)^{s-t} B_{s-1} \quad (6)$$

Sustainability implies that the second term on the RHS of (6) converges to zero as time approaches infinity. This is also known as the transversality condition which constraints the debt ratio not to grow at a faster rate than the interest rate⁷. If this is the case, then the current stock of debt should be equal to a total of both current and future discounted primary surpluses. As pointed out by Afonso (2005), the absence of no Ponzi condition can be tested empirically by testing the stock of debt for stationarity. Earlier studies that focused on testing the stationarity of public debt includes Kremers (1988), Wilcox (1989), Trehan and Walsh (1991) and Greiner and Semmler (1995).

Additionally, sustainability can be examined by testing the cointegration between revenues and expenditures, an idea initially pioneered by Hakkio and Rush (1991) and later Quintos (1995). Mathematically this can be shown from (5), by making use of the auxiliary definition $GG_{t-1} = G_t + \rho B_{t-1}$. Assuming stationary real interest rate and applying the difference operator, the present value budget constraint can be re-written as

$$GG_{t-1} - R_{t-1} = \sum_{s=t+1}^{\infty} \left(\frac{1}{1+\rho} \right)^{s-t} (\Delta R_{s-1} - \Delta G_{s-1}^r) + \lim_{s \rightarrow \infty} \left(\frac{1}{1+\rho} \right)^{s-t} \Delta B_{s-1} \quad (7)$$

Testing for the sustainability hypothesis can be done in two ways. First one could test for the absence of the no Ponzi scheme which implies that the second term of (7) approaches zero as time approaches infinity. Alternatively we could assume the absence of no Ponzi scheme and test (7) directly. In this paper, we proceed to test the absence of no Ponzi scheme.

$$\lim_{s \rightarrow \infty} \left(\frac{1}{1+\rho} \right)^{s-t} \Delta B_{s-1} = 0 \quad (8)$$

For (8) to hold, one-period government debt (ΔB_{s-1}) must not grow faster than the interest rate on debts. In other words, it is easier to see that (8) holds if ΔB_{s-1} is stationary as compared to a situation where it is not stationary. Considering that one period debt is given by the relationship $\Delta B_t = GG_t - R_t$, testing for stationarity of a

⁷Also known as the no Ponzi scheme, we rule out the possibility of the government issuing new debts in order to fund principal repayment and interest on existing debts.

one-period government debt implies testing for the difference stationarity for GG_t and R_t . This can be problematic if government spending and revenue are not stationary at their levels. However, if one can prove that they are stationary at their first difference, then the concept of cointegration can be applied. The intuition is that, if one variable can be written as a linear combination of the other with a slope coefficient such that the residual is proved to be stationary, then their relationship is stable and mean reverting. In other words, the difference of these variables does not drift wide apart. Hence we say they are cointegrated because they have a long-run stable relationship. From (8), this implies testing if GG_{t-1} and R_{t-1} are integrated of order 1 ($I(1)$) with an imposition of cointegration vector (1, -1) as argued by Quintos (1995). One can test for cointegration equivalently as below:

$$R_t = \alpha + \gamma GG_t + \mu_t \quad (9)$$

Alternatively making use of the expression $\Delta B_t = GG_t - R_t$, then from (9) we have

$$\Delta B_t = (1 - \gamma)GG_t - \alpha - \mu_t \quad (10)$$

Furthermore, $\gamma = 1$ implies sustainability since from (10) debt to GDP ratio is bounded and will grow at a constant rate. However, this condition was relaxed by Hakkio and Rush (1991), who demonstrated that the condition $0 < \gamma \leq 1$ guarantees sustainability if variables are cointegrated. Quintos (1995) argued further that $0 < \gamma \leq 1$ is both a necessary and a sufficient condition. She stressed that cointegration is only a sufficient condition for the sustainability hypothesis to hold.

At this point, it is important to make special remarks about the condition $0 < \gamma < 1$. Even though this is enough for sustainability, actually at this point the government expenditure exceeds its revenue and therefore the probability of default is high. It will be difficult to market its bonds and the government may have to pay high interest rates in order to issue new debt or attract new investors. Scenario $\gamma > 0$ guarantees sustainability since at this point, revenues are growing at a faster pace as compared to expenditures. Conversely at $\gamma < 0$, expenditures and revenues are moving in opposite directions and hence sustainability hypothesis is rejected. As shown by Quintos (1995), $\gamma = 1$ implies strong sustainability whilst γ less than 1 implies some weaker form of sustainability. Therefore the magnitude and sign of γ plays a major role in determining if indeed the sustainability hypothesis holds and the strength of the hypothesis.

Empirical cointegration test for (9) can be conducted conventionally by regressing R_t on GG_t simply by OLS and testing the residuals for stationarity in order to confirm if cointegration holds. Westerlund and Prohl (2010) argued that such conventional test fails to reject the null hypothesis of no cointegration very often, which implies a rejection of sustainability hypothesis. They cited the problem of low power of the cointegration test because of low sample size. Panel datasets circumvent the power problem as it gives an opportunity to increase the sample size. Firstly, panels present more informative data

because it has long sample size, provide more variability, less collinearity among variables and give more degree of freedom for the model (Baltagi, 2008). Secondly, panel data affords researchers the opportunity to construct and test more advanced and complicated models as compared to time series or cross-sectional data, and finally panel helps to control for the effects of omitted variables bias in econometric (Hsiao, 2003). Hence the study will resort to a panel test which will subject the residuals in (9) to a cointegration test. The test is dynamic enough to account for structural breaks and cross-sectional dependence which is common to panel data analysis. The next subsections will provide discussions on the econometric procedures for the panel test.

3 Empirical estimations and results

Firstly, we present a review of past empirical papers on fiscal sustainability with focus specifically on panel datasets. Subsequently, we will discuss our datasets and some characteristics of the data after which we shall proceed with the empirical test of the fiscal sustainability hypothesis. Table 1 shows previous papers on panel data fiscal sustainability for mostly CEEC, Organisation for Economic Corporation and Development (OECD), and EU countries. Regarding CEEC, previous studies notably by Llorca and Redzepagic (2007), Krajewski et al. (2016) and Bökemeier and Stoian (2016), all point in the direction of a sustainable fiscal policy. It will therefore be interesting to compare our results directly to these studies.

Regarding our dataset, revenue, expenditure and debt variables were all obtained from the OECD website for 10 CEEC⁸. All data exists in annual frequency. The sample period is from 1995 to 2019 and also chosen based on the availability of data. A total of 250 observations is generated from a combination of 10 countries over a 25 year period. It is important to mention that we consider total expenditures, total revenues and total debts as ratios of GDP.

Figure 1 and 2 provide a graphical overview of revenue and expenditures as well as government debt for each of the countries in the panel. We noticed that in almost all of the cases, revenues and expenditure move in the same direction even though expenditures seems to be higher than revenue for most of the time periods. Poland, Hungary and Romania displayed high variability in the revenue-expenditure relationship. Moreover, the debt to GDP ratios of Hungary and Poland for most of the years exceed revenue and expenditures. For almost all the countries, we notice a rising public debt after 2008 which can be attributed to the activeness of fiscal policy within and after the financial crises. One can infer that since spending exceeded revenue, governments borrowed more to fund their increased spending. Figure 3 shows a scatter plot that reveals the relationship between revenues and expenditure with a smooth trend line. A positive upward-sloping

⁸These countries were chosen based on the availability of quality datasets and length of time series.

Table 1: Summary of existing empirical panel fiscal sustainability test

Authors	Sustainability Test	Period and Country	Findings
Afonso and Rault (2010)	Stationarity of debt and cointegration between revenue and expenditure	15 Selected EU countries (1970 - 2006)	Fiscal stance sustainability confirmed
Baldi and Staehr(2015)	Estimated fiscal reaction function of primary balance, debt and business cycle variables	Different groups of EU countries (2001 - 2004)	Sustainable fiscal stance for all groups post financial crises
Beqiraj et. al (2018)	Panel cointegration test between primary balance and public debt	21 OECD countries (1991 - 2015)	Fiscal stance judged to be unsustainable
Bökemeier and Stoian (2016)	Fiscal reaction function of primary balance and debt	CEEC (1997 - 2013)	Fiscal stance sustainable for selected countries
Brady and Magazzino (2018)	Stationarity of public debt	19 European countries (1970 - 2016)	Fiscal stance sustainability confirmed
Checherita-Wesphal and Vaclav (2017)	Fiscal reaction function of primary balance response to debt	18 Euro Area countries (1970 - 2013)	Sustainable fiscal stance
Claeys (2007)	Cointegration between revenue, spending and net interest payment	Selected European countries (1970 - 2001)	Sustainable fiscal policy
Krajewski et al (2016)	Cointegration between revenue and expenditure and a fiscal reaction function	Central and Eastern European countries (1990 - 2012)	Sustainable fiscal stance
Lee et al. (2018)	Fiscal reaction function of primary balance response to debt	EU regional groups(1950 - 2014)	Varied results depending on the region
Llorca and Redzepagic (2007)	Cointegration between revenue and expenditure	CEEC (1999:1 - 2006:1)	Fiscal stance sustainable in selected countries
Prohl and Schneider (2006)	Cointegration between budget deficit and public debt	15 EU countries (1970 - 2004)	Fiscal stance sustainability confirmed
Westerlund and Prohl (2010)	Cointegration between revenue and expenditure	8 rich OECD countries (1977:1 - 2006:4)	Sustainability hypothesis confirmed for selected countries

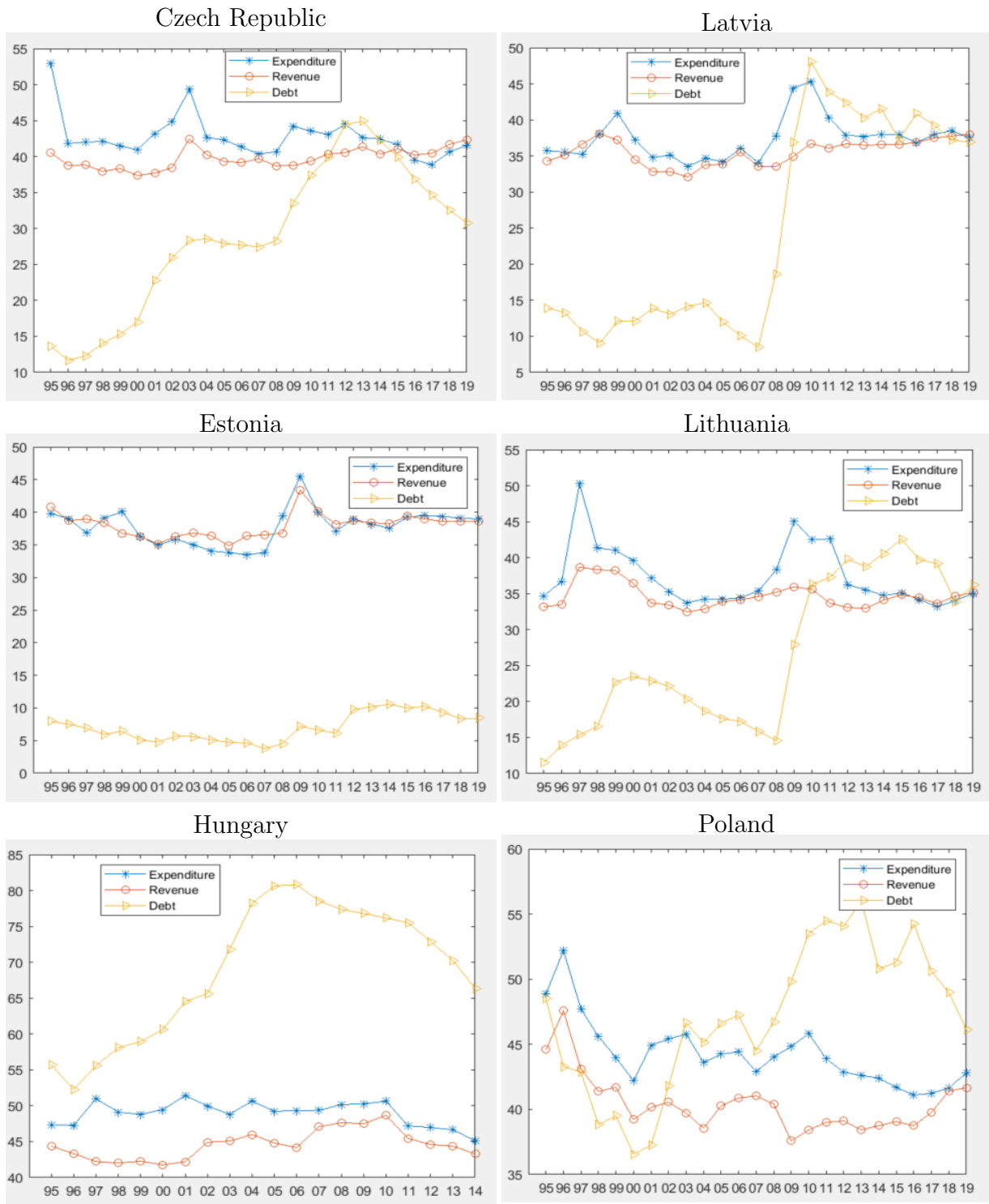


Figure 1: Revenue, Expenditure and public debt



Figure 2: Revenue, Expenditure and public debt

relationship can be observed between the two variables which provide some hints as to the nature of the relationship between the two fiscal variables.

Figure 1 and 2 also provide some hints about the possibilities of structural breaks in the individual time series. Hence it is feasible to test for the presence of structural breaks in the data. The presence of structural breaks could render statistical inferences erroneous if not accounted for in the data generating process. For instance, standard unit root tests are likely to exhibit biases toward non-rejection of the null hypothesis hence leading to a wrong conclusion about results of the test (Carrion-i-Silvestre, et al., 2005). The issue of structural break has therefore received considerable attention in both theoretical and empirical econometric literature notable among them includes Andrews et al (1996), Andrews (1993), Bai and Perron (1998) among others. Structural breaks in the mean of data and the changes in the coefficient of a linear regression coincides with

Table 2: Dates of structural breaks for individual series

Countries (CEEC)	Revenue	Expenditure
Czechia	2002,2010	2003
Estonia	1998,2008,2011	1999,2007,2010
Hungary	1997,2006,2011,2015	1998,2001,2015
Latvia	1999,2009	1997,2000,2008,2011
Lithuania	2000	2000,2008,2011
Poland	1997,2008,2015	1997,2011
Slovakia	1997,2000,2003,2012	2002,2008
Slovenia	2011,2015	2008,2015
Bulgaria	1997,2008,2012	1998,2002
Romania	1997	1998,2001,2006,2012

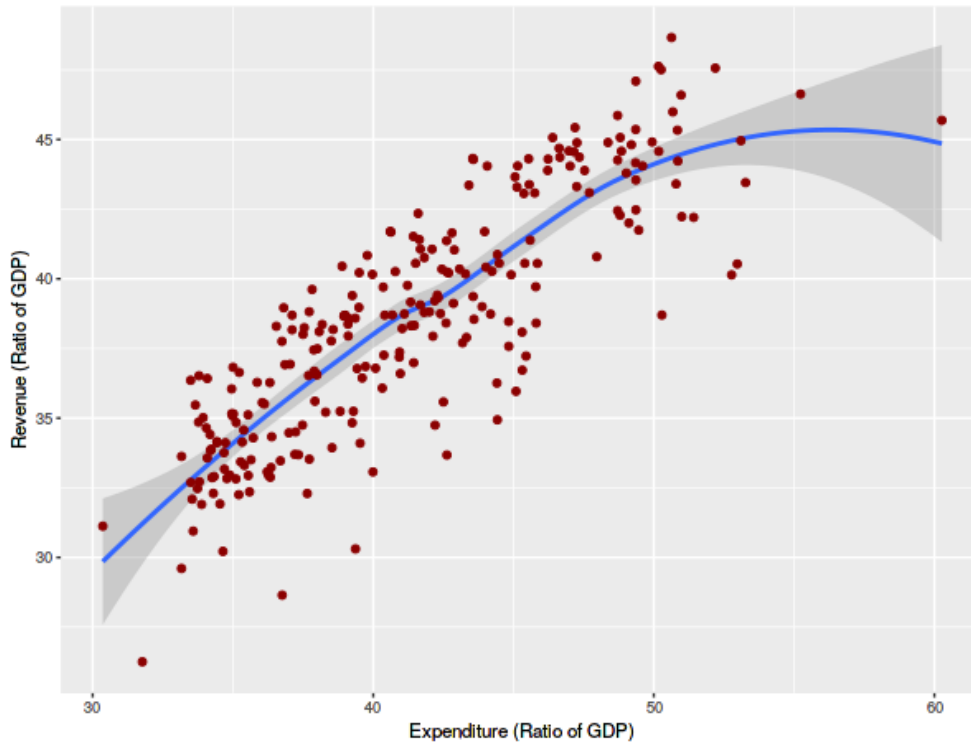
political, historical and economic events (Zeileis et al., 2003) and are therefore not usually a random phenomenon.

To test the availability of structural breaks in the individual series, this study adopts the approach by Zeileis et al (2003). There they combined the F-statistics test by Andrews (1993) and Andrews and Ploberger (1994) to test the possibility of structural breaks in regression and the technique by Bai and Perron (2003) to locate the break dates and optimal breaks in the individual series of the data ⁹. Table 2 provides results of the break dates for both revenue and expenditures whilst graphical views are displayed in Figure 5 and 6 in the appendix. Regarding revenues, the number of breaks ranges between 1 to 4. We noticed that majority of the breaks were recorded before the early 2000s which could possibly represent a policy shift as most of the CEEC were preparing to join the EU and therefore had to adjust their fiscal policies in order to meet the demands of the SPG and MT. Secondly, another break can be observed between 2007 and 2011 for most of the countries, which could also be attributed to the exogenous shock and the consequences from the global financial crises. This provides justification for the presence of the shocks and the fact that it has to be accounted for in the data generating process.

Table 3 below provides a summary statistics of the panel dataset. We notice that there is more variability in expenditures as compared to the revenue components (from the standard deviation) over the sample period. Secondly, on the average, we observe that expenditures are higher than revenues which is not so surprising since the role of government (spending) has become important especially in the twenty-first century either

⁹Procedure is implemented in R studios with the package "strucchange". We select the optimal number of breaks by choosing the number of breaks with the least sum of square residuals. Details of the structural break procedure can be found in section 5.3 in the appendix

Figure 3: Panel scatter plot



to stimulate economic growth or in direct response to macroeconomic shocks.

Table 3: Panel Summary Statistics

	Revenue-GDP ratio	Exp-GDP ratio	Debt-GDP ratio
Mean	0.391	0.419	0.340
Std. Dev.	0.039	0.505	0.178
Maximum	0.482	0.541	0.711
Minimum	0.308	0.321	0.038
Observations	250	250	250

As per the SGP requirements, member states of the EU are supposed to maintain a strict upper limit of 3% deficit to GDP ratio (Wickens, 2008), hence we investigate if member countries have followed this rule. Table 3 provides an overview of the deficit to GDP ratio of the CEEC between the sample period. We noticed that, with the exception

of Estonia that violated the SGP only once (1999), all other countries violated this rule a couple of times. Firstly it is observed that this occurred mostly between 1995 and 1998 which is prior to their accession to the EU. Secondly, during the financial crises area between 2008 up till 2012, we also notice another round of SGP violation by all countries with the exception of Estonia. CEEC have therefore run fiscal deficits over the years and have not followed the 3% deficit limit rule strictly.

Table 4: Deficit to GDP ratio

Year	Czh	Est	Hun	Lat	Lith	Pol	Svk	Slvn	Bulg	Rom
1995	-12.44	1.05	-8.60	-1.43	-1.53	-4.26	-3.47	-8.15	-5.52	-2.00
1996	-3.01	-0.32	-4.38	-0.42	-3.22	-4.63	-9.82	-1.09	-8.11	-3.57
1997	-3.19	2.15	-5.55	1.42	-11.59	-4.61	-6.27	-2.31	0.76	-4.43
1998	-4.19	-0.73	-7.39	0.03	-3.03	-4.21	-5.30	-2.33	1.08	-3.24
1999	-3.14	-3.29	-5.23	-3.74	-2.82	-2.28	-7.17	-2.97	0.09	-4.42
2000	-3.57	-0.04	-2.980	-2.73	-3.17	-2.98	-12.63	-3.65	-0.53	-4.60
2001	-5.48	0.20	-3.94	-1.95	-3.51	-4.77	-7.22	-4.45	1.05	-3.46
2002	-6.36	0.42	-8.76	-2.29	-1.85	-4.85	-8.22	-2.37	-1.16	-1.93
2003	-6.89	1.82	-7.11	-1.46	-1.27	-6.08	-3.12	-2.56	-0.39	-1.43
2004	-2.39	2.34	-6.52	-0.92	-1.39	-5.04	-2.32	-1.94	1.80	-1.09
2005	-2.9	1.08	-7.72	-0.36	-0.34	-3.96	-2.87	-1.32	1.00	-0.81
2006	-2.17	2.87	-9.21	-0.49	-0.27	-3.56	-3.58	-1.23	1.81	-2.14
2007	-0.65	2.73	-5.03	-0.51	-0.81	-1.85	-2.05	-0.05	1.10	-2.73
2008	-1.98	-2.65	-3.73	-4.20	-3.09	-3.60	-2.52	-1.39	1.59	-5.35
2009	-5.45	-2.16	-4.69	-9.49	-9.13	-7.25	-8.15	-5.81	-4.05	-9.06
2010	-4.19	0.19	-4.39	-8.60	-6.92	-7.40	-7.46	-5.60	-3.13	-6.92
2011	-2.73	1.06	-5.19	-4.25	-8.95	-4.88	-4.46	-6.63	-1.98	-5.43
2012	-3.93	-0.29	-2.27	-1.22	-3.15	-3.74	-4.37	-4.0	-0.32	-3.65
2013	-1.25	0.18	-2.54	-1.17	-2.61	-4.18	-2.87	-14.58	-0.43	-2.10
2014	-2.10	0.70	-2.76	-1.44	-0.62	-3.65	-3.11	-5.51	-5.43	-1.19
2015	-0.61	0.14	-1.97	-1.36	-0.27	-2.62	-2.67	-2.85	-1.72	-0.61
2016	0.72	-0.52	-1.76	0.06	0.23	-2.37	-2.48	-1.94	0.09	-2.62
2017	1.56	-0.77	-2.38	-0.52	0.45	-1.46	-0.95	-0.01	1.10	-2.64
2018	1.09	-0.56	-2.29	-0.74	0.60	-0.24	-1.06	0.77	1.75	-2.96
2019	0.75	-0.30	-1.83	0.51	0.13	-1.16	-1.03	0.72	-0.11	-3.33

Highlights in bold indicates violation of the EU Stability and Growth Pact. Source: author's own computations

Next, we test for evidence of cross-sectional dependence of the individual units in the

panel. We deem it feasible to test for cross-sectional dependence which is peculiar with macro panel data because countries in the same region respond to shocks in the similar ways, hence generating serially correlated errors which affect inferences from the econometric test. Previous first generational econometric unit root test such as Levin, Lin and Chu (2002), Im, Pesaran and Shin test(2003) and Maddala and Wu (1999) and cointegration test notably Pedroni (2000, 2004) and Kao (1999) assume that the cross-section in the panel data is independent. Such test usually suffer from size distortions which affects the inferences (Banerjee et al, 2004). Properly accounting for cross-sectional dependence in panels improves the efficiency of parameter estimates and simplifies statistical inferences (Hsiao, 2014). Two main cross-sectional dependence test namely Breusch-Pagan test and Pesaran test are carried out in this paper. Proposed by Breusch and Pagan (1980), the test is based on a Lagrangian multiplier which is applicable to heterogeneous models and other variant panel models. Breusch-Pagan test is very convenient for datasets with short N and large T (Pesaran, 2004). Further details about the cross-sectional dependence test can be found in section 5.2 in the appendix.

Table 5 presents results of the Breusch-Pagan and Pesaran cross-sectional dependence test. The null hypothesis for both tests indicates cross sectional independence in the panel datasets. In the case of expenditures to GDP ratio, there is a strong rejection of the null hypothesis for both tests. Hence we accept the alternative hypothesis of cross-sectional dependence. For revenue to GDP ratio, there is a strong rejection (1%) for the Breusch-Pagan test and rejection at 10% significance level for the Pesaran test. Result provides evidence of cross country dependence in the panel data and hence justifies the need to choose an econometric procedure that accounts for cross-sectional dependence.

Table 5: Breusch-Pagan and Pesaran Cross sectional dependence Test

Variables	Breusch-Pagan CD Test		Pesaran CD Test	
	chisq	p-value	z-value	p-value
Expenditure to GDP ratio	107.84	0.000	5.494	0.000
Revenue to GDP ratio	113.42	0.000	1.716	0.0862

Pooled group variable by country. Null hypothesis of the test implies Cross sectional independence for both Breusch-Pagan and Pesaran test.

From an econometric point of view, it is important to decide if indeed data can be pooled or not. According to Baltagi et al. (2008), imposing the pooling restriction reduces the variance of the pooled estimator. However, this could lead to a bias and hence wrong inferences if the restriction is false. Pooling data is based on the assumption that the parameters in the model are the same (homogeneous) across the individual countries. Similarly, we can also verify if the parameters are the same or different across the time periods. The decision of whether to pool or not is a natural question which arises in

Table 6: Chow test of poolability of coefficients (5% significance level)

Restricted Model	Intercept	F-Statistics	P-value	Verdict
FE(Within)	Variable	2.8438	0.004	Not Poolable
Pooled OLS	Fixed	14.892	0.000	Not Poolable

Null hypothesis implies "model stability" or constant coefficient

panel studies (Baltagi, 2005). Once the true nature of the parameters is established, it is then feasible to choose an appropriate econometric estimator. Considering an unrestricted model in the regression of the form

$$R_{it} = \alpha_i + \gamma_j GG_{it} + \epsilon_{it} \quad (11)$$

Where R_{it} is government revenue to GDP ratio, G_{it} is government expenditure to GDP ratio, ϵ_{it} represents the error term or residuals and α_i is the time-invariant intercept. From the slope coefficient, j could be the individual countries (heterogeneous across the countries) or time (heterogeneous across the time period). In each case, we use the Chow (1960) test analogous to the F-test to test for poolability under the assumption that the residuals are normally distributed with a zero mean and constant variance. The test statistics is constructed by looking at the difference between the Sum of Squared Residuals (SSR) of the restricted model and the SSR of the unrestricted model and dividing by the SSR of the unrestricted model considering their degrees of freedom. In the context of panels, the unrestricted model could be a Fixed Effect (FE) within the model (with variable intercepts) or a pooled OLS regression model (constant intercept). A detailed discussion of the Chow test can be found in Wooldridge(2009) and Baltagi (2005).

From Table 6, if we consider an unrestricted model of a FE-within model with a time-invariant slope, then clearly we can reject the null of model stability or poolability at 1% significance level. Similarly, if we consider a pooled OLS regression with constant intercept and slope parameter, we can still reject the null hypothesis at 5% level. Hence there is evidence to back the claim that we cannot pool the slope coefficient across the individual countries from our data. The slope coefficient is therefore heterogeneous across countries and cannot be considered as homogeneous. This provides useful guidance regarding the selection of an appropriate econometric model and procedure suitable for accounting for heterogeneity in the slope coefficient.

As part of the cointegration requirement, the variables have to be integrated of order 1. In other words, we test if revenues and expenditures are stationary at their first difference ($I(1)$). In this study, we adopt the Fourier unit root test by Nazlioglu and Karul (2017) which allows for smooth breaks in the mean of the series and cross-sectional dependence at the same time. This test is one of the few second generational unit root tests that accounts for both cross-sectional dependence and structural breaks. The test is a combination of

an earlier test by Becker et al. (2006) who employed a Fourier approximation function to model structural breaks and Hadri and Kazoumi (2011, 2012) who used a common factor structure to account for cross-sectional dependence. A Fourier approximation can be used to model structural shifts of any form or non-linearity in the deterministic term as this was shown by Becker et al (2006). It is important to note that the Fourier approximation is used to model breaks or shifts in a smooth gradual process which is in contrast to sharp breaks¹⁰. Another distinctive feature of the test is that the breaks are determined endogenously and do not have to be pre-determined. A detailed discussion of the test can be found in section 5.4 in the appendix.

The null hypothesis of the test implies "stationarity" against the alternative of a unit root. The test depends on the Fourier frequency (k) which determines the swings and amplitude of the series. Considering the span of the time series in the panel (25 years), we choose $k = 3$, sufficient enough to cover the length of the time series. Tables 6 and 7 present the results of the panel univariate stationarity test and the test for the individual countries for expenditure and revenue respectively¹¹. We observe that the null of stationarity for the panel is strongly rejected at 1% significance level irrespective of whether we consider a model with a "constant" or a model with "constant and a trend". For the individual countries, as k increases, we fail to reject the null of stationarity for most of the countries. However, considering the panel test statistic, the variables have a unit root and are therefore not stationary. It is necessary to test the first difference to ensure they are $I(1)$. Stationarity test for the first difference of the variables shows the absence of unit root. Table 13 and 14 in the appendix (for the sake of space) show the stationarity test results of the first difference of revenue and expenditures. It is observed that there is lack of evidence to reject the null hypothesis completely when we consider a model with a "constant" and in some cases a "constant and a trend". As a robustness check, we employ the unit root test by Carrion-i-Silvestre et al. (2005) which accounts for structural breaks to test for if indeed the variables are $I(1)$. Results in Table 15 in the appendix support the claim that revenue and expenditures are $I(1)$. We therefore conclude that revenues and expenditure are stationary at their first difference.

After establishing that revenue and expenditure are $I(1)$, we estimate the cointegration relationship between the variables. Regarding testing the cointegration relationship between revenues and expenditure, we resort to the test by Westerlund and Edgerton (2008). This test is very appealing because it serves as a one-stop-shop by accounting for structural breaks and cross-sectional dependence in panel data making it very desirable. Secondly, the test is robust to serial correlation and heteroscedasticity in the residuals. Westerlund and Edgerton (2008) proposed two tests for the null hypothesis of no cointe-

¹⁰Models with sharp breaks cannot be modelled by Fourier approximation. In such instances, dummy variables can be used to capture the sharp breaks. Carrion-i-Silvestre et al. (2005) developed a panel unit root test which is capable of accommodating sharp breaks in panels by using dummy variables.

¹¹I would like to thank Saban Nazlioglu for making the Gauss codes available.

Table 7: Panel stationarity Test - Expenditure

Countries	Constant			Constant and trend		
	$k = 1$	$k = 2$	$k = 3$	$k = 1$	$k = 2$	$k = 3$
Czechia	0.070	0.309	0.302	0.050*	0.059	0.050
Estonia	0.099	0.410*	0.331	0.052*	0.124*	0.125*
Hungary.	0.203**	0.298	0.201	0.044	0.101	0.091
Latvia	0.183**	0.504**	0.414*	0.051*	0.080	0.083
Lithuania	0.052	0.055	0.116	0.051*	0.043	0.051
Poland	0.182**	0.496*	0.416*	0.053*	0.056	0.076
Slovakia	0.048	0.219	0.197	0.040	0.145**	0.139*
Slovenia	0.059	0.189	0.322	0.054*	0.097	0.085
Bulgaria	0.148**	0.089	0.095	0.051*	0.088	0.087
Romania	0.124	0.313	0.181	0.053*	0.136**	0.133*
Panel statistic	2.995*** (0.001)	3.508*** (0.000)	2.282*** (0.011)	4.938*** (0.000)	3.309*** (0.000)	2.460*** (0.007)

Fourier Panel Stationarity Test for 10 CEEC under the Null hypothesis of Stationarity. P-values are for one sided test based on normal distribution.

Critical values (obtained from Becker et al. 2006:289) for individual test statistics are as follows 0.1318 (10%), 0.1720 (5%), 0.2699 (1%) for $k=1$; 0.3150 (10%), 0.4152 (5%), 0.6671 (1%) for $k=2$; 0.3393 (10%), 0.4480 (5%), 0.7182 (1%) for $k=3$

Critical values for constant and trend are as follows 0.0471 (10%), 0.0546 (5%), 0.0716 (1%) for $k=1$; 0.1034 (10%), 0.1321 (5%), 0.2022 (1%) for $k=2$; 0.1141 (10%), 0.1423 (5%), 0.2103 (1%) for $k=3$

Significance at 10%, 5% and 1% are denoted by *, ** and *** respectively.

gration. The proposed test is derived from a Lagrangian multiplier (LM) function in the similitude of Schmidt and Phillips (1992), Ahn (1993), and Amsler and Lee (1995) unit root based test. Thorough details of the cointegration test can be found in section 5.5 in the appendix.

We test the null of "no cointegration" against an alternative hypothesis of "cointegration" between revenue and expenditures. There are two proposed test statistics of the null hypothesis. The first test statistics $Z_{\tau}(N)$ is based on the least square estimate of the residual slope whilst the second test statistics $Z_{\phi}(N)$ is based on estimating the t-ratio of the slope. A maximum of 3 breaks is chosen for the cointegration relationship. The selection of optimum lag length is based on an automatic procedure adopted from Campbell and Perron (1991).

Regarding the output of the test, we consider three scenarios. Firstly, we test the null hypothesis of "no cointegration" under the condition of absence of breaks. That is, we assume there are no breaks in the cointegration relationship. Secondly, we test the null

Table 8: Panel stationarity Test - Revenue

Countries	Constant			Constant and trend		
	$k = 1$	$k = 2$	$k = 3$	$k = 1$	$k = 2$	$k = 3$
Czechia	0.146*	0.314	0.408*	0.047	0.048	0.068
Estonia	0.157*	0.271	0.108	0.044	0.066	0.091
Hungary.	0.342***	0.187	0.148	0.040	0.107*	0.103
Latvia	0.070	0.448**	0.313	0.052*	0.074	0.076
Lithuania	0.074	0.155	0.299	0.051*	0.076	0.054
Poland	0.170*	0.409*	0.352*	0.065**	0.101	0.103
Slovakia	0.262**	0.371*	0.362*	0.057**	0.153**	0.143**
Slovenia	0.056	0.128	0.239	0.056*	0.079	0.074
Bulgaria	0.268**	0.115	0.121	0.065**	0.114*	0.102
Romania	0.074	0.151	0.160	0.049*	0.121*	0.116*
Panel statistic	5.652*** (0.000)	2.716*** (0.003)	2.136*** (0.016)	5.606*** (0.000)	3.405*** (0.000)	2.523*** (0.006)

Fourier Panel Stationarity Test for 10 CEEC under the Null hypothesis of Stationarity. P-values are for one sided test based on normal distribution.

Critical values (obtained from Becker et al. 2006:289) for individual test statistics are as follows 0.1318 (10%), 0.1720 (5%), 0.2699 (1%) for $k=1$; 0.3150 (10%), 0.4152 (5%), 0.6671 (1%) for $k=2$; 0.3393 (10%), 0.4480 (5%), 0.7182 (1%) for $k=3$

Critical values for constant and trend are as follows 0.0471 (10%), 0.0546 (5%), 0.0716 (1%) for $k=1$; 0.1034 (10%), 0.1321 (5%), 0.2022 (1%) for $k=2$; 0.1141 (10%), 0.1423 (5%), 0.2103 (1%) for $k=3$

Significance at 10%, 5% and 1% are denoted by *, ** and *** respectively.

hypothesis by considering breaks in only the intercept (level break). Finally, we consider breaks in both the intercept and the slope (regime shift). From Table 9 none of the models is cointegrated when we consider significance at a strict 5% level. Considering a more relaxed significance level at 10%, we find evidence of cointegration for the model with no breaks for the $Z_\tau(N)$ test and no cointegration for $Z_\phi(N)$. Hence even with no breaks, the cointegration relationship is not strongly confirmed. This provides fresh evidence of lack of cointegration between total revenues and expenditures (all ratios of GDP)¹². This implies a rejection of the fiscal sustainability hypothesis for CEEC which is in contrast to previous studies on CEEC notably by Krajewski et. al (2016) and Llorca and Redzepagic (2007). Even though they both employed a panel cointegration procedure, their studies did not test for structural breaks and cross-sectional dependence in the cointegration relationships which can be considered a major weakness. Hence accounting for this dynamism (breaks and cross-sectional dependence) in a panel data setting reinforces the credibility of the

¹²I would like to thank Joakim Westerlund for making the Gauss codes available.

results in this study.

Table 9: Panel Cointegration test of Revenue and Expenditure - Europe

Models	$Z_\tau(N)$		$Z_\phi(N)$	
	Value(τ)	P-Value	Value(ϕ)	P-value
No breaks	-1.515	0.065	-0.963	0.168
Level break	0.879	0.810	0.690	0.755
Regime shift	0.422	0.663	0.437	0.669
Number of observations	250		250	

Westerlund and Edgerton(2008) cointegration test with three maximum number of breaks in the cointegration relationship which are determined by grid search at the minimum of the sum of squared residuals. Null hypothesis indicates "No cointegration". Displayed p-values is based on one-sided normal distribution test. *, ** and *** denotes rejection of the null hypothesis at 10%, 5% and 1% respectively..

3.1 Adjusting fiscal variables for cyclicity

Recall from (9) the cointegration relationship between revenues and expenditure. We decompose these fiscal variables into a trend and cyclical components. Following Gali et al. (2003), we posit that the cyclical component of fiscal variables necessitates automatic responses from government which represents passive policy. In other words, this aspect does not constitute planned long-term government action and is influenced mainly by business cycles. The trend component on the other hand represents an active discretionary fiscal policy and hence should be only considered when examining the long-term behaviour of government policy. Decomposing (9), we have

$$R_t^c + R_t^\tau = \alpha + \gamma(GG_t^c + GG_t^\tau) + \mu_t \quad (12)$$

Where R_t^c and R_t^τ represent cyclical and trend components of revenue whilst GG_t^c and GG_t^τ are cyclical and trend components of government expenditures respectively. Statistically, the cyclical component of variables is mean reverting hence stationary. In other words, the cyclical component represents short-run dynamics which will eventually die out in the long run. Secondly, in a panel cointegration set up, stationary and zero mean variables will end up in the residual term of (12) and will therefore not influence the cointegrating vector, hence it is justifiable from an econometric perspective to exclude the cyclical component in the cointegration relationship (Beqiraj et al, 2018). Therefore from (12), we end up with

$$R_t^\tau = \alpha + \gamma(GG_t^\tau) + \mu_t \quad (13)$$

A popular tool for decomposing a series into trend and cyclical component is the Hodrick Prescott(HP) filter (see Hodrick and Prescott (1997)). Consider a time series of the form

$$y_t = \tau_t + c \quad (14)$$

Using the Hodrick Prescott filter, we denote mathematically by minimizing the equation

$$Min_{\tau} \left(\sum_{t=1}^T (y_t - \tau_t)^2 + \lambda \sum_{t=1}^T [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right) \quad (15)$$

Where y_t denotes the actual series at period t , τ_t denotes the trend component at time t and c represents the cyclical component of the series. λ denotes the smoothing parameter which is key to estimating the trend such that as λ approaches 0, the trend component approaches the actual time series whereas λ approaches infinity implies a linear trend. Empirical value of $\lambda=1600$ was used by Hodrick and Prescott (1997) for US quarterly data. However for annual data $\lambda = 100$ is recommended (Martin et al., 2013).

In a seminal paper, Hamilton (2018) proposed an alternative method for de-trending a series and proved that the HP filter is deficient in three respects. Firstly he argued that HP filter imposes a spurious dynamic relationship which has no basis as far as the data generating process is concerned. Secondly, there are discrepancies between filtered values at the end of the sample and those at the middle of the sample and also spurious. Finally, the values of HP smoothing parameter is vastly at odds with common practice and hence not reliable. To demonstrate his recommended approach, Hamilton (2018) applied OLS regression of series y_t on a constant and four recent values of y at time t as

$$y_{t+h} = \beta_0 + \beta_1 y_t + \beta_2 y_{t-1} + \beta_3 y_{t-2} + \beta_4 y_{t-3} + u_{t+h} \quad (16)$$

Where y_t represents a quarterly time series and h is a 8 quarter time horizon which is approximately two years¹³. The residual u_t which is assumed stationary represents the cyclical component of the original series y_t and is given by

$$\hat{u}_{t+h} = y_{t+h} - \hat{\beta}_0 - \hat{\beta}_1 y_t - \hat{\beta}_2 y_{t-1} - \hat{\beta}_3 y_{t-2} - \hat{\beta}_4 y_{t-3} \quad (17)$$

The residual is stationary provided the fourth difference of y_t is stationary. The study will adopt both filters in order to ascertain if results after using filters differ significantly.

As a requirement for cointegration, we test all variables at their levels and first difference to ensure they are $I(1)$. Results in the appendix indicate that cyclically adjusted revenue and expenditure have a unit root in their levels and are stationary in first difference paving the way for the cointegration test. Table 10 provides result of the cointegration test between cyclically adjusted revenues and cyclically adjusted expenditure.

¹³In the case of data with annual frequency, Hamilton(2018) recommended $h = 2$.

When we consider a model with no breaks, we rejected the null at a strict 1% level which implies accepting the alternative hypothesis of cointegration. Secondly, there is strong cointegration when we consider breaks in the level or regime shift with a strong rejection of the null hypothesis of no cointegration at 1% level for both $Z_\tau(N)$ and $Z_\phi(N)$. The result is not any different if we use cyclically-adjusted variables from the HP filter (See Table 16 in the appendix). The result strongly supports cointegration between cyclically adjusted revenue and expenditures for CEEC countries. This implies that if we consider cyclically adjusted variables, we can infer that government in CEEC have jointly pursued a sustainable fiscal policy.

Table 10: Panel Cointegration test of cyclically adjusted revenue and cyclically adjusted spending

Models	$Z_\tau(N)$		$Z_\phi(N)$	
	Value(τ)	P-Value	Value(ϕ)	P-value
No breaks	-2.842***	0.002	-3.587***	0.000
Level break	-3.076***	0.001	-2.950***	0.002
Regime shift	-2.368***	0.009	-2.984***	0.001
Number of observations	220		220	

Westerlund and Edgerton (2008) cointegration test with three maximum number of breaks in the cointegration relationship which are determined by grid search at the minimum of the sum of squared residuals. Null hypothesis indicates "No cointegration". Displayed p-values is based on one-sided normal distribution test. *, ** and *** denotes rejection of the null hypothesis at 10%, 5% and 1% respectively. Hamilton Filter is used to obtain cyclically adjusted revenues and expenditures.

3.2 Estimation of cointegration vector

Once cointegration is established, it is necessary to estimate the equilibrium parameter in the long run dynamic relationship in order to infer the sustainability hypothesis. Recall from section (2) that to infer strong sustainability in the sense of Quintos (1995), the size of the slope coefficient must be equal to unity, otherwise a weak form of stationarity is inferred. It is important to note that when variables are expressed as ratios of GDP or in per capita terms, then it is even crucial to get a slope coefficient of 1 in other to ensure that debt is bounded and does not explode to infinity (Afonso, 2005). We estimate a panel form of (13) with heterogeneous slope coefficient as below

$$R_{it}^\tau = \alpha + \gamma_i(GG_{it}^\tau) + \mu_{it} \quad (18)$$

Even though the Ordinary Least Square estimator (OLS) has been found to be super consistent, it has been proven to be deficient with finite data sets and complex dynamic relationships (Forest and Turner, 2013) and also insufficient for panel data (Baltagi, 2005).

Considering the fact that our slope coefficient in (18) is heterogeneous, it is necessary to use estimates that account for slope heterogeneity. This is very common with macro datasets with large time series and cross-sectional dimensions represented by countries or regions. The asymptotic of macro panels are different from that of traditional micro panels with large number of cross-sections and a smaller number of time period. Such micro datasets are usually estimated with Fixed Effects (FE), random effects, instrumental variables techniques to account for possible endogeneity such as the Generalized method of moments (Blackburne and Frank, 2007). However, these estimators rely on the assumption that the slope coefficient is homogeneous, something which is not applicable to our model.

Likely candidates are the mean group type of estimates which includes the Mean group estimator (MG), Common Correlated Effect Mean Group (CCMG) and Augmented Mean Group. The Mean group estimator first developed by Pesaran and Smith (1995) is similar to the fixed effect-within model, however, it averages the slope for each individual country in the panel. From (18), one would estimate the N-group specific ordinary least squares regression and average the estimated coefficient for the group in order to account for heterogeneity in the coefficient. From (18), MG estimator is as follows

$$\hat{\gamma}_{mg} = \frac{1}{N} \sum_{n=1}^N \hat{\gamma}_{ols,n} \quad (19)$$

One of the main criticisms of the MG estimator is the fact that it does not account for the issue of cross-sectional dependence in panel data. Hence inferences from this estimator should be made with some caution due to potential bias. To circumvent this problem, Pesaran (2006) developed the CCEMG that accounts for cross-sectional dependence by allowing for heterogeneous impact across panel members. From (18), we expand the error term to include an unobserved common factor which is recovered by cross sectional averages of the dependent variable and independent variable $\mu_{it} = \lambda F_{it} + \eta_{it}$.

Where F_{it} is the common factor term and η is a random shock. Once the unobserved common factor is recovered, the estimator of the group can be obtained by again averaging the slope coefficients across the panel in a similar fashion as the MG estimator. This estimator is therefore robust against cross-sectional dependence (see Pesaran (2006) for further details).

A further problem arises if the regressor in the model is potentially endogenous. Then most estimators which do not account for endogeneity bias will suffer from depending on nuisance parameter (Westerlund and Prohl, 2010). Authors such as Kao and Chiang (2000) and Chen, McCosky and Kao (1999) therefore recommended the Fully modified OLS (FMOLS) proposed by Phillips and Hansen (1990) and Dynamic OLS (DOL) introduced by Saikkonen (1990) and later advanced by Stock and Watson(1993) as promising models for estimating the long-run vector in a cointegration regression.

Due to the fact that our slope coefficient is very heterogeneous, MG estimator of FMOLS and DOLS will be appropriate in this context. Specifically, the MG-FMOLS and MG-DOLS introduced by Pedroni (2000, 2001) are suitable for estimating the cointegration vector such that it is consistent with cross-sectional heterogeneity in panel cointegration studies. Again consider a panel regression of the form in (18)

$$R_{it} = \alpha_i + GG'_{it}\gamma_i + u_{it}$$

Where R_{it} is the dependent variable (1x1) and γ is the vector of slope parameter, α_i represents the intercept. $i = 1, \dots, N$ and $t = 1, \dots, T$. u_t is the disturbance term which is assumed to be stationary. GG_{it} is a $k \times 1$ regressor vector which is assumed to follow the process

$$GG_{it} = GG_{it-1} + \epsilon_{it}$$

Under the specification above, if R_{it} and GG_{it} are assumed to be cointegrated, they are both integrated process of order 1 (written for notational simplicity as $I(1)$). The OLS estimator for (42) is given by

$$\hat{\gamma}_{OLS} = \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(GG_{it} - \bar{G}G_i)' \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T (GG_{it} - \bar{G}G_i)(R_{it} - \bar{R}_i) \right] \quad (20)$$

The FMOLS estimator makes correction for the OLS model by accounting for endogeneity and serial correlation in the OLS estimator in (18) by applying a non parametric correction. The MG - FMOLS estimator (accounting for heterogeneity in slope coefficient) is given by

$$\hat{\gamma}_{MG-FMOLS} = \left[N^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T (GG_{it} - \bar{G}G_i)(GG_{it} - \bar{G}G_i)' \right]^{-1} * \left[\sum_{i=1}^N \left(\sum_{t=1}^T (x_{it} - \bar{G}G_i) \hat{R}_{it}^+ - T \hat{\Delta}_{\epsilon u}^+ \right) \right] \right] \quad (21)$$

Where $\hat{R}_{it}^+ = \alpha_i + GG'_{it}\gamma_i + u_{it} - \hat{\Omega}_{u\epsilon}\hat{\Omega}_{\epsilon}^{-1}\Delta GG_{it}$ is the endogeneity correction term, such that $\hat{\Omega}_{\epsilon}$ and $\hat{\Omega}_{u\epsilon}$ are consistent estimates of Ω_{ϵ} and $\Omega_{u\epsilon}$ respectively. Where Ω is the covariance matrix of GG and R and $\hat{\Delta}_{\epsilon u}^+$ is the serial correlation correction term given by $\hat{\Delta}_{\epsilon u} - \hat{\Delta}_{\epsilon}\hat{\Omega}_{\epsilon}^{-1}\hat{\Omega}_{\epsilon u}$.

The MG-DOLS regression on the other hand entails augmenting the cointegration model with lags and leads of ΔGG_{it} so that there is orthogonality between the error term

and the regressors. This corrects the endogeneity and serial correlation in the panel cointegration regression and the concept of the mean group is to account for the heterogeneity across cross sections. The MG-DOLS estimator $\hat{\gamma}_{DOLS}$ is obtained from the equation

$$R_{it} = \alpha_i + GG'_{it}\gamma_i + \sum_{j=-q}^q c_{ij}\Delta GG_{it+j} + \bar{v} \quad (22)$$

Where \bar{v} is the combination of the disturbance terms given by $\bar{v} = v_{it} + \sum_{j>q} c_{ij}\epsilon_{it+j}$.

From (22), the panel DOL estimator is given as

$$\hat{\gamma}_{MG-DOLS} = \left[N^{-1} \sum_{i=1}^N \left(\sum_{t=1}^T z_{it}z'_{it} \right)^{-1} \left(\sum_{t=1}^T z_{it}\tilde{s}_{it} \right) \right] \quad (23)$$

Where $z_{it} = GG_{it} - \bar{G}G_{it}, \Delta GG_{it-k}, \dots, \Delta GG_{it+k}$ is a vector of regressors and $\tilde{s}_{it} = R_{it} - \bar{R}_i$.

The study makes use of all four models (MG, CCEMG, MG-FMOLS and MG-DOLS) to estimate the long-run cointegration vectors. Table 11 shows the estimated cointegration slope (γ) using the different estimators. In the case of MG-DOLS, we explore lags and leads from 1 to 4 to experiment with the sensitivity of the γ coefficient. For FMOLS, we make use of the Bartlett Kernel for the the long-run covariance matrix. The long-run coefficient for the MG and CCEMG are 0.499 and 0.364 respectively and p-values indicate their statistical significance with low standard errors. When we consider the MG-FMOLS and MG-DOLS, the long-run coefficients are 0.938 and 0.935 respectively which are higher comparatively. The probability values indicate that the estimated slope is statistically significant with low relative standard errors.

We conduct some residual diagnostic tests to ascertain which estimator performs better. Recall that the model (18) relies on the assumption that the residuals are normally distributed with a zero-mean and a constant variance. Hence it is feasible to test if indeed this is the case. The lower part of Table 11 depicts test statistics and probability values of the Shapiro-Wilk test (see Royston (1982)). From the p-values, we can reject the null hypothesis of "normality in residuals" for the MG and CCEMG estimators at 1% significant level. In the case of MG-FMOLS and MG-DOLS, we cannot reject the null hypothesis of residual normality, hence these two models perform better because their residuals are normally distributed. Figure 4 shows a graphical view of the normality of the residuals for all the estimators by way of quantile-quantile plot. It is observed that there is an almost perfect correlation between the normal distribution theoretical quantile and the sample quantile of MG-DOLS and MG-FMOLS estimates. Hence this goes to reinforce the results of the Shapiro-Wilk test that the residuals of these two estimators are normally distributed. Regarding residuals from MG and CCEMG estimators, we notice the residuals do not fall on the straight line and hence confirms the results of the Shapiro-Wilk test results of residual non-normality.

Table 11: Long run coefficient for cyclically adjusted fiscal variables

Stat	MG	CCEMG	MG-FMOLS	MG-DOLS ^a
γ	0.499	0.364	0.938	0.935
Test stat	3.390	3.815	480.74	604.07
P-val	0.000	0.000	0.000	0.000
Std error	0.147	0.095	0.002	0.02
Obs	220	220	210	130
Shapiro-Wilk Normality test	0.970	0.977	0.992	0.986
P-val	(0.000)	(0.001)	(0.260)	(0.220)
Peseran CD Test	3.652	-2.411	3.328	-0.260
P-val	(0.000)	(0.016)	(0.001)	(0.795)

a - Reported lags and leads of 4 for MG-DOLS. The study explored lags and leads from 1 to 4, however this does not change the estimates of the parameter.

Secondly, we conduct cross-sectional dependence test making use of the Pesaran test (discussed in section 5.2 in the appendix). Reported p-values from table 11 reveal that the null hypothesis of cross sectional independence can be rejected for MG, CCEMG and MG-FMOLS at 1%, 5% and 1% respectively in favour of the alternative hypothesis of the presence of cross-sectional dependence in the residuals. In the case of MG-DOLS model, we cannot reject the null hypothesis even if we consider a lax 10% significance level. This proves that the MG-DOLS is robust against cross-sectional dependence and has normally distributed residuals making it the most efficient estimator among the three.

Considering the size of the slope coefficient for the two efficient estimators (MG-DOLS and MG-FMOLS), it is important to establish if indeed they are equal to 1. Recall from section 2 that a slope coefficient of 1 guarantees strong fiscal sustainability since it implies that the debt to GDP ratio is bounded. To ascertain if cointegration slope (γ) is indeed 1, it is plausible to conduct a hypothesis test of the coefficient. We employ the Wald test under the null hypothesis that $(\gamma) = 1$ ($H_0 : \gamma = 1$), as against the alternative that $H_a : \gamma < 1$. The Wald test statistics takes the form:

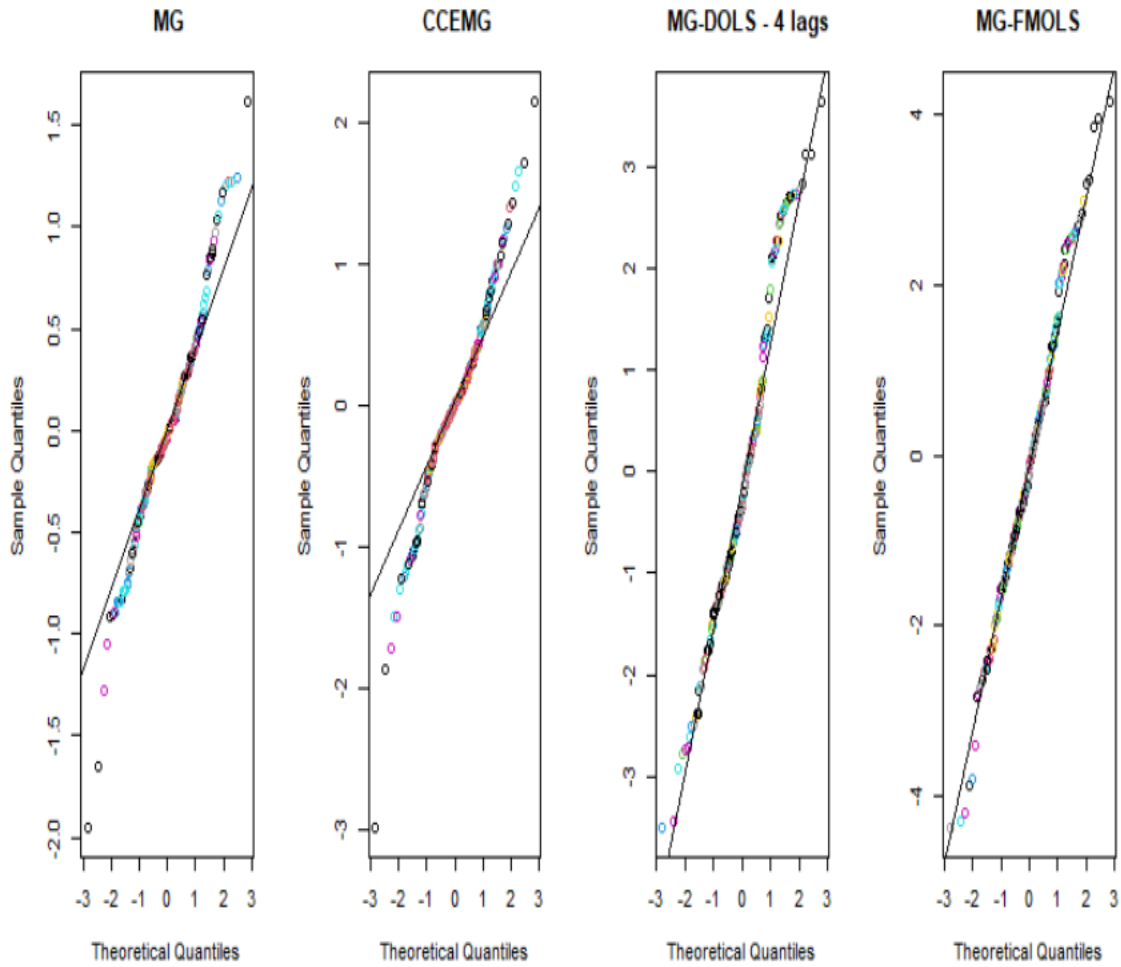
$$W = (\hat{\gamma} - \gamma_0)'[var(\hat{\gamma})]^{-1}(\hat{\gamma} - \gamma_0) \sim \chi_p^2 \quad (24)$$

Which reduces to

$$W = \frac{(\hat{\gamma} - \gamma_0)^2}{var(\hat{\gamma})} \sim \chi_p^2 \quad (25)$$

Where $\hat{\gamma}$ is the maximum likelihood estimate of the parameter to be tested, γ_0 is the parameter which is assumed to be true under the null hypothesis. If the null hypothesis

Figure 4: Residual diagnostics: Quantile-Quantile plot



is true then W is chi-square distributed with p degrees of freedom which also represents the number of parameters to be estimated.

Results of the Wald test shown in Table 12 imply the rejection of the null hypothesis of a unit slope for the two models (MG-DOLS and MG-FMOLS) indicating that $0 < \gamma < 1$. This is statistically significant if we consider the p-values of the Wald-test. Further, we construct confidence intervals to show the position of the true value of γ at 95% level. All evidence shows that $\gamma < 1$ which implies weak sustainability for cyclically adjusted variables in the sense of Quintos (1995). Even though cyclically adjusted revenue and expenditure are cointegrated, the magnitude of the cointegration slope is not strong enough to guarantee strong sustainability. The intuition is that considering a linear

Table 12: Wald test of Coefficient and confidence intervals

Models	MG-FMOLS	MG-DOLS
T stat	-32.018***	-49.89***
Chi-Square (1 df)	1025.13	1754.72
P-value	0.000	0.000
95% confidence intervals	(0.934 - 0.941)	(0.932 - 0.938)

*, ** and *** denotes rejection of the null at 10%, 5% and 1% respectively. Null hypothesis: $\gamma = 1$.

equation (18), an increase in expenditure by 1 unit will induce revenue to increase by less than 1 unit all other things being equal. Hence expenditure to GDP ratio grows more than revenue to GDP implying the accumulation of debts and hence a bubble debt term in the long run. Even though there is cointegration for cyclically adjusted variables, debt to GDP ratio is not finite in the long run. Hence we refer to the fiscal stance of CEEC as weakly sustainable in the sense of Qunitos (1995). Based on the above findings, we therefore conclude that CEEC jointly have pursued a weakly sustainable fiscal policies if we consider cyclically adjusted revenues and spending to GDP ratio.

4 Conclusion

This study sought to ascertain if the fiscal sustainability hypothesis holds for 10 CEEC from the period 1995 to 2019. Previous studies have shown that these countries have pursued policies compatible with the government intertemporal budget constraint. We tested the hypothesis of sustainability of the fiscal stance by examining the cointegration relationship between revenues and expenditures, both as percentages of GDP. The econometric intuition is that if revenues and expenditure can be expressed as a linear combination and residuals can be proven to be stationary, then debt to GDP ratio is mean reverting since the difference between revenue and expenditures do not drift wide apart. Hence inferences about long term relationship between revenues and expenditures could be made.

We adopted recent advancements in econometrics to test the fiscal sustainability hypothesis. As a first step, we considered total revenues and total expenditure. Preliminary results indicated that these fiscal variables are not cointegrated and cast doubt on the sustainability hypothesis for the 10 CEEC. The result is also in sharp contrast to earlier panel studies conducted for CEEC which have all pointed in the direction of cointegrated revenue and expenditures. However, none of the studies considered accounted for structural breaks and cross-sectional dependence in the data generating process, something that has become associated with dynamic macro panels. The study therefore tested,

found evidence and accounted for structural breaks for CEEC - most of which occurred as a result of the changes in fiscal policies prior to joining the EU and also shocks due to business cycles notably the global financial crises in 2008.

As a next step, the study makes a justification for using cyclically adjusted revenues and expenditures and argues that this represents the long term discretionary action of the fiscal authorities. Hence the action of fiscal authorities should be judged by variables which are devoid of business cycle fluctuations or shocks. This is plausible because shocks to fiscal variables induce an automatic response by policymakers and do not necessarily characterize discretionary policy. We use the recently formulated Hamilton filter which addresses the limitations of the popular HP filter to obtain cyclically-adjusted fiscal variables. Results indicate that cyclically adjusted revenue and expenditures are cointegrated with a slope less than unity. We employed the Wald test to ascertain if indeed the slope coefficient is unity by way of hypothesis testing since the values are close enough to unity. Results provide enough evidence to reject the null hypothesis of a unit slope coefficient indicating that the coefficient lies between 0 and 1. Considering the fact that these variables are ratios to GDP, a unit slope of the cointegration is necessary to guarantee strong sustainability in the sense of Quintos(1995). But even though there is cointegration between cyclically adjusted revenue and expenditure, a slope coefficient less than unity implies that expenditures to GDP ratio will grow faster than revenues to GDP ratio implying a weaker form of sustainability. This is because the debt to GDP ratio is not bounded and therefore not finite. If this continues to happen for a long time, it will generate spikes in the debt to GDP ratio and the fiscal stance will no longer be sustainable.

The possible policy implications are as follows. Firstly, holders of government bonds could lose confidence if debt accumulation is persistent since this casts doubt with regards to the ability of the government to service its payment. Secondly, the government may have difficulties in marketing its debts to new investors and hence would not be able to raise substantial additional revenue by issuing bonds in the future due to unattractiveness of its debts. Otherwise, government would have to pay high interest in order to make its debt attractive to investors. CEEC governments may therefore have to alter their fiscal policy by way of increasing revenue or reducing expenditure or both as a way of counteracting the deficit problem. The study provides fresh evidence using cyclically adjusted revenue and expenditure for panel sustainability analysis in the context of CEEC. The discretionary action of the government is deemed not to be sufficient to infer strong sustainability of the fiscal stance. The government in CEEC must therefore do more to address the fiscal deficit problem by way of fiscal consolidation to avoid future implications of sustainability.

With the current corona pandemic, fiscal sustainability has become even more challenging as the current recession necessitates further action of the government in terms of stimulating aggregate demand. However with low revenues due to low productivity and output, government cannot respond adequately to the pandemic without for instance

borrowing to augment its revenue. Others have also advocated for taxing the super-rich in society as a way of increasing revenue. However, the effectiveness of this policy as demonstrated by Scheuer and Slemrod (2019) depends on the elasticity of the taxpayers. The current recession and the previous (global financial crises) have taught us that the possibility of a looming recession in the future cannot be ruled out, hence there should be adequate fiscal space for governments to respond appropriately to future shocks. It is therefore important for government with high debt burdens to institute structural changes, especially in normal times as a way of reducing debt stocks. This will ensure that there is enough fiscal space in the future to combat the consequences of recessions.

References

- Afonso, A. (2005). "Fiscal sustainability: the unpleasant European case." *FinanzArchive*, Vol. 61, pp. 19–44.
- Ahmed, S., & Rogers, J. (1995). "Government Budget Deficits and Trade Deficits. Are Present Value Constraints Satisfied in Long-term Data?." *Journal of Monetary Economics*, Vol. 36, pp. 351-374.
- Ahn, S. K. (1993). "Some tests for unit roots in autoregressive-integrated-moving average models with deterministic trends" *Biometrika*, Vol. 80, pp.855-868
- Amsler, C., & Lee, J. (1995). "An LM test for a unit root in the presence of a structural break" *Econometric Theory*, Vol. 11, pp.359–368.
- Andrews, D.W.K. (1993). "Test for parameter instability and structural change with unknown change point" *Econometrica*, Vol. 61, pp.821-856.
- Andrews, D.W.K., Lee, I., & Ploberger, W. (1996). "Optimal changepoint tests for normal linear regression" *Econometrica*, Vol. 70, pp.9-36.
- Andrews, D.W.K., & Ploberger, W. (1994). "Optimal test when a nuisance parameter is present only under the alternative" *Econometrica*, Vol. 62, pp.1383-1414.
- Bai, J., & Ng, S. (2004). "A panic attack on unit roots and cointegration" *Econometrica*, Vol. 72, pp.1127–1177.
- Bai, J., & Perron, P. (2003). "Computation and analysis of multiple structural change models" *Econometrica*, Vol. 66, pp.1-22.
- Bai, J., & Perron, P. (1998). "Estimating and Testing Linear Models with Multiple Structural Changes" *Econometrica*, Vol. 66, pp.47–78
- Baltagi, H.B. (2008). "Forecasting with panel data" *Journal of Forecasting*, Vol. 27(2), pp.153-173.
- Baltagi, H.B. (2005). "Econometric analysis of panel data" *Third edition, John Wiley and Sons, Chichester, England*.

- Banerjee, A., Marcellino, M., & Osbat, C. (2004). "Some cautions on the use of panel methods for integrated series of macro-economic data" *Econometrics Journal*, Vol. 7(2), pp.322-34.
- Becker, R., Enders, W., & Lee, J. (2006). "A stationarity test in the presence of an unknown number of smooth breaks" *Journal of Time Series Analysis*, Vol. 27(3), pp.381–409.
- Beqiraj, E., Fedeli,S., & Forte, F. (2018). "Public debt sustainability: An empirical study on OECD countries" *Journal of Macroeconomics*, Vol. 58, pp.238-248.
- Blackburne, E.F., & Frank, M.W. (2007). "Estimation of Nonstationarity Heterogeneous Panels". *The Stata Journal*, Vol. 7, pp. 197-208.
- Blanchard, O. (2006). "Macroeconomics, Fourth edition" *Pearson Prentice Hall, New Jersey*.
- Bökemeier, B., & Stoian, A. (2018). "Debt Sustainability Issues in Central and East European Countries." *Eastern European Economics*, Vol. 56(5), pp. 438-470.
- Breusch, T.S., & Pagan, A. (1980). "The LM Test and Its Applications to Model Specification in Econometrics" *Review of Economic Studies*, Vol. 47, pp.239-254.
- Campbell, J., & Perron, P. (1991). "Pitfalls and opportunities: what macroeconomists should know about unit roots" *NBER Macroeconomics Annual, MIT Press, Cambridge, MA*, Vol. 72, pp. 141–201.
- Carrion-i-Silvestre, J.L., Barrio-Castro, T.d., & Lopez-Bazo, E. (2005). "Breaking the panels: An application to the GDP per capita" *Econometrics Journal* , Vol. 8(2), pp.159–175
- Chen, B., McCoskey, S., & Kao, C. (1999). "Estimation and inference of a cointegrated regression inpanel data: A Monte Carlo study" *American Journal of Mathematical and Management Sciences*. Vol. 19, pp. 75–114.
- Claeys, P. (2007). "Sustainability of EU Fiscal Policies: a Panel Test" *Journal of Economic Integration*, Vol. 22(1), pp.112-127.
- Forest, J.J., & Turner, P. (2013). "Alternative estimators of cointegrating parameters in models with nonstationary data: an application to US export demand" *Applied Economics*. Vol. 45(5), pp. 629-636.
- Galí, J., Perotti, R., Lane, R.P., & Richter, F.W. (2003). "Fiscal Policy and Monetary Integration in Europe" *Econ. Policy*, Vol. 18(37), pp.535–572.
- Gleich, H. (2003). "Budget institutions and fiscal performance in Central and Eastern European countries". *European Central Bank Working Paper Series*, No. 215.
- Greiner, H., & Fincke, B. (2015). " Public Debt, Sustainability and Economic Growth." *Heidelberg: Springer Verlag*.

- Hadri, K., & Kurozumi, E. (2011). "A locally optimal test for no unit root in cross-sectionally dependent panel data" *Hitotsubashi Journal of Economics*, Vol. 52(2), pp.165-184.
- Hadri, K., & Kurozumi, E. (2012). "A simple panel stationarity test in the presence of serial correlation and a common factor" *Economics Letters*, Vol. 115(1), pp.31-34
- Hallett, A.H., & Lewis, J. (2007). "Debt, deficits, and the accession of the new member States to the Euro" *European Journal of Political Economy*, Vol. 23(2), pp.316-337.
- Hamilton, J. D. (2018). "Why you should never use the Hodrick-Prescott filter" *Review of Economics and Statistics*, Vol. 100(5), pp. 831–843.
- Hamilton, J., & Flavin, M. (1986). "On the Limitations of Government Borrowing: A Framework for Empirical Testing." *American Economic Review*, Vol. 76, pp. 808 - 819.
- Hsiao, C. (2003). "Analysis of Panel Data, Second, edition" *Cambridge University Press*, UK.
- Im, K. S., Pesaran, M. H., & Shin, Y. (2003). "Testing for unit roots in heterogeneous panels" *Journal of Econometrics*, Vol. 115, pp.53-74.
- Kao, C. (1999). "Spurious regression and residual-based test for cointegration in panel data" *Journal of Econometrics*, Vol. 90, pp.1-44.
- Kao, C., & Chiang, M.H. (2000). "On the estimation and inference of a cointegrated regression in panel data" *Advances in Econometrics*. Vol. 15, pp. 179-222.
- Krajewski, P., Mackiewicz, M., & Szymańska, A. (1993). "Fiscal Sustainability in Central and Eastern European Countries - A Post-Crisis Assessment" *Prague Economic Papers* , Vol. 25(2), pp.175-188
- Kremers, J. (1988). "The Long-Run Limits of U.S. Federal Debt." *Economics Letters*, Vol. 28, pp. 259 - 262.
- Levin, A., Lin, C., & Chu, C. (2002). "Unit root tests in panel data: Asymptotic and Finite-sample properties" *Journal of Econometrics*, Vol. 108, pp.1-24.
- Liu, P., & Tanner, E. (1995). "Intertemporal Solvency and Breaks in the U.S. Deficit Process: A Maximum-likelihood Cointegration Approach." *Economics Letters*, Vol. 2, pp. 231-235.
- Maddala, G.S., & Wu, S. (1999). "A comparative study of unit root test with panel data and a new simple test" *Oxford Bulletin of Economics and Statistics*, Vol. 61, pp.631-652.
- Martin, V., Hurn, S., & Harris, D. (2013). "Econometric Modelling with Time Series : Specification, Estimation and Testing" *Themes in Modern Econometrics*, Cambridge University Press, New York

- Nazlioglu, S., & Karul, C. (2017). "A panel stationarity test with gradual structural shifts: re-investigate the international commodity price shocks" *Economic Modelling*, Vol. 61, pp.181-192.
- OECD. (2000). "Government at a Glance - yearly updates, Public finance and economics" Available at <https://stats.oecd.org/Index.aspx?queryid=82342>, Accessed: 19 September 2020.
- Pedroni, P. (2004). "Panel cointegration: asymptotic and finite sample properties of pooled time series test with an application to the PPP Hypothesis" *Econometric Theory*, Vol. 61, pp.597-625.
- Pedroni, P. (2001). "Purchasing power parity tests in cointegrated panels" *Review of Economics and Statistics*. Vol. 83, pp.727-731.
- Pedroni, P. (2000). "Fully modified OLS for heterogeneous cointegrated panels" *Advances in Econometrics*, Vol. 15, pp.93-130.
- Pesaran, M. H. (2007). "A simple panel unit root test in the presence of cross-section dependence" *Journal of Applied Econometrics*, Vol. 22(2), pp.265-312
- Pesaran, M.H. (2004). "General diagnostics test for cross-sectional dependence in panels", *Working Paper, Trinity College, Dublin*.
- Pesaran, M. H., & Smith, R. (1995). "Estimating long-run relationships from dynamic heterogeneous panels," of *Econometrics*, vol. 68(1), pp. 79-113.
- Phillips, P.C.B., & B.E. Hansen. (1990). "Statistical inference in instrumental variables regression with I(1) processes" *Review of Economic Studies*. Vol. 57, pp. 99-125.
- Prohl, S., & Schneider, F. (2006). "Sustainability of public debt and budget deficit: Panel cointegration analysis for the European Union Member countries " *Working Paper, No. 0610, Johannes Kepler University of Linz, Department of Economics, Linz*.
- Quintos, C. (1995). "Sustainability of the Deficit Process with Structural Shifts." *Applied Journal of Business Economic Statistics*, Vol. 13, pp. 409 -417.
- Royston, P (1982). "An extension of Shapiro and Wilk's W test for normality to large samples" *Applied Statistics*, Vol. 31, pp. 115-124.
- Saikkonen, P. (1991). "Asymptotically Efficient Estimation of Cointegrating Regressions" *Econometric Theory*. Vol. 7, pp.1-21.
- Scheuer, F. & Slemrod, J. (2019). "Taxation and the Superrich", *NBER Working Paper Series*, No. 26207.
- Schmidt, P., & Phillips, P. C. B. (1992). "LM tests for a unit root in the presence of deterministic trends", *Oxford Bulletin of Economics and Statistics*, Vol. 54, pp.257-287.

- Stock J., & M. Watson, (1993). "A simple estimator of cointegrating vectors in higher order integrated systems" *Econometrica*. Vol. 61, pp. 783–820.
- Trehan, B. & Walsh, C. (1988). "Common Trends, the Government's Budget Constraint, and Revenue Smoothing." *Journal of Economic Dynamics and Control*, Vol. 12, pp. 425 - 444.
- Westerlund, W., & Edgerton, D.L (2008). "A Simple Test for cointegration in Dependent Panels with Structural Breaks" *Oxford Bulletin of Economic and Statistics*, Vol. 70(5), pp.665–704.
- Westerlund, J., & Prohl, S. (2010). "Panel cointegration tests of the sustainability hypothesis in rich OECD countries." *Appl. Econ.*, Vol. 42, No.1, pp. 1355-1364.
- Wickens, M. (2008). "Macroeconomic theory,a dynamic general equilibrium approach" *Princeton University Press*.
- Wilcox, D., & Walsh, C. (1989). "The Sustainability of Government Deficits: Implications of the Present- Value Borrowing Constraint." *Journal of Money, Credit, and Banking* , Vol. 21, pp. 291-306.
- Wooldridge, J.M. (2009). "Introductory econometrics: a modern approach." *Fourth edition* , South-Western.
- Zeileis, A., Kleiber, C., Krämer, W., & Hornik, K (2003). "Testing and dating structural dates in practice" *Computational Statistics and Data Analysis*, Vol. 44, pp.109-123.

5 Appendix

5.1 Proof of Equation (6) by forward substitution

Making B_{t-1} the subject from (5) ,

$$B_{t-1} = \frac{R_t}{(1+\rho)} - \frac{G_t}{(1+\rho)} + \frac{B_t}{(1+\rho)} \quad (26)$$

Then we can write

$$B_{t-2} = \frac{R_{t-1}}{(1+\rho)} - \frac{G_{t-1}}{(1+\rho)} + \frac{B_{t-1}}{(1+\rho)} \quad (27)$$

Substituting (50) into (51) and re-arranging terms, we get

$$B_{t-2} = \frac{R_{t-1}}{(1+\rho)} + \frac{R_t}{(1+\rho)^2} - \left(\frac{G_{t-1}}{(1+\rho)} + \frac{G_t}{(1+\rho)^2} \right) + \frac{B_t}{(1+\rho)^2} \quad (28)$$

From (50), we can write

$$B_{t-3} = \frac{R_{t-2}}{(1+\rho)} - \frac{G_{t-2}}{(1+\rho)} + \frac{B_{t-2}}{(1+\rho)} \quad (29)$$

Substituting (52) into (53) and re-arranging terms, we get

$$B_{t-3} = \frac{R_{t-2}}{(1+\rho)} + \frac{R_{t-1}}{(1+\rho)^2} + \frac{R_t}{(1+\rho)^3} - \left(\frac{G_{t-2}}{(1+\rho)} + \frac{G_{t-1}}{(1+\rho)^2} + \frac{G_t}{(1+\rho)^3} \right) + \frac{B_t}{(1+\rho)^3} \quad (30)$$

We can continue to derive equation for B_{t-4}

$$B_{t-4} = \frac{R_{t-3}}{(1+\rho)} + \frac{R_{t-2}}{(1+\rho)^2} + \frac{R_{t-1}}{(1+\rho)^3} + \frac{R_t}{(1+\rho)^4} - \left(\frac{G_{t-3}}{(1+\rho)} + \frac{G_{t-2}}{(1+\rho)^2} + \frac{G_{t-1}}{(1+\rho)^3} + \frac{G_t}{(1+\rho)^4} \right) + \frac{B_t}{(1+\rho)^4} \quad (31)$$

Doing this recursively up to B_{t-n} and summing from $s = t + 1$ to some k , for all $s, k > 0$ we get

$$B_{t-n} = \sum_{s=t+1}^k \left(\frac{1}{1+\rho} \right)^{s-t} R_{s-n} - \sum_{s=t+1}^k \left(\frac{1}{1+\rho} \right)^{s-t} G_{s-n} + \lim_{s \rightarrow \infty} \left(\frac{1}{1+\rho} \right)^{s-t} B_{s-n} \quad (32)$$

Hence summing from $s = t + 1$ to infinity, B_{t-1} can be written as

$$B_{t-1} = \sum_{s=t+1}^{\infty} \left(\frac{1}{1+\rho} \right)^{s-t} (R_{s-1} - G_{s-1}) + \lim_{s \rightarrow \infty} \left(\frac{1}{1+\rho} \right)^{s-t} B_{s-1} \quad (33)$$

□

5.2 Cross sectional dependence Test

Consider a panel model of the form

$$y_{it} = \alpha_i + \beta_i x_{it} + \epsilon_{it} \quad (34)$$

Where the individual cross sections are represented by $i = 1, \dots, N$, the time period captured by $t = 1, \dots, T$, x represents the vector of explanatory variables, α_i is a time invariant parameter and parameter β_i varies across the panels. The LM statistics derived

by Breusch and Pagan is as follows:

$$LM_{CD} = T \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij}^2$$

$$\text{Where } \hat{\rho}_{ij} = \frac{\sum_{t=1}^T e_{it}e_{jt}}{(\sum_{t=1}^T e_{it}^2)^{1/2}(\sum_{t=1}^T e_{jt}^2)^{1/2}}$$

$\hat{\rho}_{ij}$ is the sample estimator of the pairwise correlation of the residuals where e_{it} represents the OLS estimate of the residual given by $e_{it} = y_{it} - \hat{\alpha}_i - \hat{\beta}_i' x_{it}$.

The parameters $\hat{\alpha}_i$ and $\hat{\beta}_i$ are estimates of α_i and β_i which are obtained after regressing y_{it} on x_{it} and its intercept for each of the cross sections. The null hypothesis of the Breusch-Pagan test indicates "no cross sectional dependence in the panel". For a panel data sets with N larger than T , the Breusch-Pagan test tends to over-reject the null hypothesis (Hsiao, 2014). This may therefore be problematic for such panels. Pesaran (2004) proposed a way of circumventing this problem of finite sample bias for micro datasets with large N and small T .

To address the bias emanating from the adjustment of finite sample of the Breusch-Pagan LM test and its related complications, Pesaran (2004) proposed a test statistics for the cross sectional dependence (CD) test. The test statistics of the Pesaran test is given by

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij} \right)$$

$$\text{where } \hat{\rho}_{ij} = \frac{\sum_{t=1}^T e_{it}e_{jt}}{(\sum_{t=1}^T e_{it}^2)^{1/2}(\sum_{t=1}^T e_{jt}^2)^{1/2}}$$

Once again $\hat{\rho}_{ij}$ denotes the pairwise correlation of the OLS residuals. Pesaran (2004) showed that the test has satisfactory power in small samples and is robust to unit roots and structural breaks

5.3 Testing and dating of Structural breaks

Consider a model of the form

$$y_t = x_t^\tau \beta_i + u_t \quad (35)$$

Where y_t is the observation of dependent variable, x_t is the explanatory variable which is $k \times 1$ vector, β_i is a $k \times 1$ vector as well and u_t is the disturbance term. Following Zeileis et al (2003), we test the null hypothesis that the coefficients in (35) are constant against the alternative hypothesis of at least one non-constant coefficient in the regression.

$$H_0 : \beta_i = \beta_0 \quad (36)$$

Where $i = 1, \dots, n$. If we assume m number of break points, that is to say shifts of regression coefficient m times, then there will be $m+1$ segments with constant coefficients. Hence (36) becomes

$$H_0 : \beta_j = \beta_0 \quad (37)$$

Where $i = i_{j-1} + 1, \dots, i_j$, j (Segment index) = $1, \dots, m + 1$. Using the F-statistics, we test the null hypothesis of (37) at $m = 1$ which is a single shift of alternative timing. The test is based on the F-statistics for a change at time t , with one regression for each sub-sample. Each break point is compared to the residual \hat{u} in the below model,

$$F_i = \frac{\hat{u}^\tau \hat{u} - \hat{u}(i)^\tau \hat{u}(i)}{\hat{u}(i)^\tau \hat{u}(i) / (n - 2k)} \quad (38)$$

where $\hat{u}(i)$ is the estimate of $u(i)$ which represents the segmented residuals in case of breaks. From (38), the F-statistics is computed for $i = n_h, \dots, n - n_h$ ($n_h \geq k$) and the rejection of the null hypothesis is based on the size of the supremum or average. If the average computed F-statistics is too large, we reject the null. Recall from (14) that the null indicates "absence of a structural break". The computation of the asymptotic p-value of the test is based on Hansen (1997).

Next will be to locate the number of breaks and exact break dates which in practice is rarely given exogeneously. Suppose an m -partition i_1, \dots, i_m , then the least square estimates for β_j can be obtained and the resulting minimal residual sum of squares is

$$RSS(i_1, \dots, i_m) = \sum_{j=1}^{m+1} r_{ss}(i_{j-1} + 1, i_j) \quad (39)$$

Where $r_{ss}(i_{j-1} + 1, i_j)$ is the minimal SSR in the j th segment. Hence we seek to find the breakpoints $\hat{i}_1, \dots, \hat{i}_m$ which will minimize the objective function

$$(\hat{i}_1, \dots, \hat{i}_m) = \operatorname{argmin} RSS(i_1, \dots, i_m) \quad (40)$$

With partitions (i_1, \dots, i_m) and $i_j - i_{j-1} \geq n_h \geq k$. To proceed, we seek to obtain the global minimizers of (40) by a grid search, but this will be computationally lot of work if done manually especially if the number of breaks $m > 2$. This can simply be obtained by adopting Bai and Perron (2003), who developed a dynamic programming procedure for the structural change model in the context of OLS regression. Based on the Bellman principle, their approach entails finding an optimal segmentation that satisfies the recursive equation

$$RSS(\mathfrak{k}_{m,n}) = \min[(\mathfrak{k}_{m-1,i}) + r_{ss}(i + 1, n)] \quad (41)$$

Where $\mathfrak{k}_{m,n}$ denotes the break point set or the m -partition and $i_0 = 0$ and i_{m+1} . Further details regarding the dynamic programming procedure for the location of the break can be seen in Bai and Perron (2003).

5.4 Panel Unit root test

Consider a data generating process such as the below

$$y_{it} = \alpha_i(t) + x_{it} + \lambda F_t + \epsilon_{it} \quad (42)$$

$$x_{it} = x_{it-1} + u_{it} \quad (43)$$

Where $i = 1, \dots, N$ represents the cross sections and $t = 1, \dots, T$ is the time dimension. $\alpha_i(t)$ is the deterministic time variant term and x_{it} is a random walk process having initial values x_{i0} for all i . The disturbance terms ϵ_{it} and u_{it} are mutually independent as well as independent and identically distributed (i.i.d) across the cross sections and over the time dimension with

$$E(\epsilon_{it}) = 0, E(\epsilon_{it}^2) = \sigma_{\epsilon_i}^2 > 0, E(u_{it}) = 0, E(u_{it}^2) = \sigma_{u_i}^2 > 0$$

F_t represents the unobserved common factor and λ_t is the weight or factor loadings. Where ϵ_{it} , F_t and λ_t are all i.i.d whilst F_t is assumed to be stationary and serially uncorrelated with $E(F_t) = 0$ and $E(F_t^2) = \sigma_F^2 > 0$. The intercept including an unknown form of the structural shift is modelled in accordance with the Fourier approximation as the below

$$\alpha_i(t) = a_i + \rho_{1i} \sin\left(\frac{2\pi kt}{T}\right) + \rho_{2i} \cos\left(\frac{2\pi kt}{T}\right) \quad (44)$$

Alternatively, to capture the breaks in the intercepts and slope, the deterministic trend is modelled as

$$\alpha_i(t) = a_i + b_i(t) + \rho_{1i} \sin\left(\frac{2\pi kt}{T}\right) + \rho_{2i} \cos\left(\frac{2\pi kt}{T}\right) \quad (45)$$

Where ρ_{1i} and ρ_{2i} measures the amplitude(wavelength) and displacement of the shifts, k denotes the Fourier frequency which determines the distribution of the individual test statistic. The null hypothesis implies stationarity whilst the alternative hypothesis specifies a unit root. The alternative hypothesis allows some stationarity in the cross section and permits $\sigma_{u_i}^2$ to vary across cross sections as well.

The individual test statistic is based on KPSS test allowing Fourier frequency according to Becker et. al (2006). The test statistic is written as

$$TS(k) = \frac{1}{T^2} \frac{\sum_{t=1}^T \tilde{S}_{it}(k^2)}{\sigma_{\epsilon_i}^2}$$

Where $\tilde{S}_{it}(k) = \sum_{j=1}^t \tilde{\epsilon}_j$ is the partial sum emanating from the OLS residual and $\sigma_{\epsilon_i}^2$ is the long run variance estimate.

To obtain the Fourier panel statistics, the individual test statistics is averaged to obtain

$$FP(K) = \frac{1}{N} \sum_{i=1}^N TS(k).$$

The asymptotic distribution of the Fourier test statistics as $T \rightarrow \infty$ depends only on the Fourier frequency k and does not depend on other parameters. To deal with the issue of cross sectional dependence in panels, Nazioglu and Karul (2017) followed Pesaran (2007) by taking cross sectional average of the model and replacing the common factor with the estimate of the average of the cross section. To demonstrate this, let $Z_t = [1, \sin(2\pi kt/T), \cos(2\pi kt/T)]'$ or $Z_t = [1, t, \sin(2\pi kt/T), \cos(2\pi kt/T)]'$. Then the cross sectional average is given by

$$\bar{y}_t = Z_t' \bar{\sigma} + \bar{r}_t + \bar{\lambda} F_t + \bar{\epsilon} \quad (46)$$

By assuming $\lambda \neq 0$ and $N \rightarrow \infty$, the solution of (23) is $F = \bar{y} - Z_t' \bar{\sigma} - \bar{r}_t - \bar{\epsilon} / \bar{\lambda}$. Using this result leads to

$$y_{it} = Z_t' \tilde{\sigma}_i + \tilde{\lambda}_i \bar{y}_t + e_{it} \quad (47)$$

Where $\tilde{\sigma}_i = \sigma_i - \bar{\sigma} \tilde{\lambda}_i$, $\tilde{\lambda}_i = \lambda_i / \bar{\lambda}$, $e_{it} = r_{it} \bar{r}_t + \epsilon_{it} - \tilde{\lambda}_i \bar{\epsilon}_t$ and \bar{r}_t and $\bar{\epsilon}_t$ goes to infinity under the null hypothesis. Hence the common factor can be replaced by the average of the cross sections of $y_{it}(\bar{y}_t)$. This allows the regression of y_{it} on Z_t and \bar{y}_t for each of the cross sections.

Nazioglu and Karul applied monte carlo simulations to show that the test has good size and power properties even in small sample provided the errors are i.i.d. Secondly, they proved that in the face of serial correlation of errors the test still has reasonable size and high power.

5.5 Panel Cointegration Test

Consider a variate y_{it} which evolves according to :

$$y_{it} = \alpha_i + \eta_i t + \delta D_{it} + x_{it}' \beta_i + (D_{it} x_{it})' \gamma_i + z_{it} \quad (48)$$

$$x_{it} = x_{it-1} + w_{it} \quad (49)$$

Where x_{it} is a k dimensional vector of regressors modelled as a random walk process, w_{it} is the error term, $i = 1, \dots, N$ represents the cross sectional dimension, $t = 1, \dots, T$ represents the time period. D_{it} is a scalar representing a break dummy such that $D_{it} = 1$ if $t > T_i$ and zero otherwise. Parameters α_i and β_i are the intercept and slope respectively before the break, whilst σ_i and γ_i are the changes in these parameters after the shift ¹⁴.

Next we assume the disturbance term z_{it} is modelled in order to allow for cross sectional dependence by way of unobserved common factor as :

¹⁴For the sake of notational simplicity, the model is restricted to only one break. However, generalizations can be made to accommodate multiple breaks.

$$z_{it} = \lambda_i' F_t + v_{it} \quad (50)$$

$$F_{jt} = \rho_j F_{jt-1} + u_{jt} \quad (51)$$

$$\phi_i(L) \Delta v_{it} = \phi_i v_{it-1} + e_{it} \quad (52)$$

where $\phi_i(L) := 1 - \sum_{j=1}^{p_i} \phi_{ij} L^j$ represents a scalar polynomial in the lag vector L , F_t represents a vector of unobservable common factor F_{jt} where $j = 1, \dots, r$ and lastly λ_t represents vector of factor loading. In order to impose strict stationary, we assume that $\rho_j < 1$ for all j . This then implies from (50) that the order of integration of z_{it} is dependent on the idiosyncratic disturbance term v_{it} . We can infer then that the data generating relation in (48) is cointegrated provided $\phi < 0$ and spurious if $\phi = 0$. The null hypothesis of the test is that all N cross sectional units are spurious ($H_0 : N_1 = 0$) whilst the alternative hypothesis states that the first N_1 units are cointegrated with the remaining $N_0 := N - N_1$ considered spurious ($H_1 : N_1 > 0$)¹⁵.

Testing for the null as against the alternative hypothesis, Westerlund and Edgerton (2008) employed a Lagrangian Multiplier (LM) principle which states that the score vector has a zero mean when estimated at the vector close to the true parameters under the null. They considered a pooled log-likelihood function of the form:

$$\log(L) = constant - \frac{1}{2} \sum_{i=1}^N (T \log(\sigma_i^2) - \frac{1}{\sigma_i^2} \sum_{t=1}^T e_{it}^2)$$

The test is then derived by concentrating the log-likelihood function w.r.t σ_i^2 and thereby evaluating the resulting score at the restricted maximum likelihood function. Assuming $\hat{\sigma}_i^2 := \frac{1}{T} \sum_{t=1}^T e_{it}^2$, then each unit's score contribution is given by

$$\frac{\partial \log(L)}{\partial \phi_i} = \frac{1}{\hat{\sigma}_i^2} \sum_{t=2}^T (\Delta \hat{S}_{it} - \Delta \hat{S}_i) (\Delta \hat{S}_{it-1} - \Delta S_i)$$

where \hat{S}_{it} is a residual which is defined as

$$\Delta \hat{S}_{it} = constant + \phi_i \hat{S}_{it-1} + error \quad (53)$$

$\Delta \hat{S}_i$ and \hat{S}_i represent the mean values of $\Delta \hat{S}_{it}$ and \hat{S}_{it-1} respectively. From (53) the score vector is proportional to the numerator of the least squares estimates of ϕ_i . Next, we can proceed to formulate the test of the null hypothesis of no cointegration for the cross sectional unit i as a zero slope restriction of (53). This can be done in two ways, firstly one can test the least squares estimate of ϕ_i , secondly by testing the t-ratios. A

¹⁵Again for notational simplicity it is assumed that the first N units are cointegrated under the alternative hypothesis. This is therefore by no means a restriction.

panel test (of H_0 versus H_1) can therefore be constructed by compiling a cross sectional sum of the individual i statistics making use of the log-likelihood functional form. The presence of cross sectional dependence is modelled as

$$\hat{S}_{it} = y_{it} - \hat{\alpha}_i - \hat{\sigma}_i D_{it} - \hat{\eta}_i t - x'_{it} \hat{\beta}_i - (D_{it} x_{it})' \gamma_i - \hat{\lambda}_i \hat{F}_t.$$

Where \hat{F}_t is an unobserved common factor which is obtained by summing the principal component estimate $\Delta \hat{F}$ of ΔF . Hence defactoring makes the test robust against cross sectional dependence typically emanating from the common factors in the cross sections. Also by including the lag of the dependent variable, we make the test robust against serial correlation as shown below:

$$\Delta \hat{S}_{it} = constant + \phi_i \hat{S}_{it-1} + \sum_{j=1}^{pi} \phi_{ij} \Delta \hat{S}_{it-j} + error \quad (54)$$

The first test statistic is defined as

$$LM_{\phi}(i) = T \hat{\phi}_i \left(\frac{\hat{\omega}_i}{\hat{\sigma}_i} \right) \quad (55)$$

Where $\hat{\phi}_i$ is the least square estimate of ϕ_i from (54), σ_i is the estimated standard error of the regression and ω_i^2 is the estimated long run variance of Δv_{it} based on \hat{S}_{it} . The second test statistics derived from the t-ratio of $\hat{\phi}_i$ is defined as

$$LM_{\tau}(i) = \frac{\hat{\phi}_i}{SE(\hat{\phi}_i)} \quad (56)$$

Where $SE(\hat{\phi}_i)$ is the estimated standard error of $\hat{\phi}_i$. It follows that, the two panel based LM test proposed for the null of no cointegration is given by

$$L\bar{M}_{\phi}(N) = \frac{1}{N} \sum_{i=1}^N LM_{\phi}(i) \quad (57)$$

$$L\bar{M}_{\tau}(N) = \frac{1}{N} \sum_{i=1}^N LM_{\tau}(i) \quad (58)$$

To estimate the unknown structural breaks, Westerlund and Edgerton (2008) followed Bai and Perron (1998) by estimating the breakpoints individually for each cross sectional unit and by minimizing the sum of square residuals from the first difference of the regression in (48). Hence the break point estimator can be defined as

$$\hat{\tau}_i = \arg \min_{0 < \tau < 1} \frac{1}{T-1} \sum_{t=2}^T (\Delta \hat{z}_{it})^2,$$

When estimating the common factors, the recommendation of Bai and Ng (2004) is followed, which assumes the minimization of an information criterion. The choice of the number of lags length p_i in (54) can be made using a data dependent rule based on the significance of the individual lag parameters ϕ_i as suggested by Campbell and Perron (1991). Alternatively an information criterion such as the Schwarz Bayesian criterion could also be used to estimate the number of lags needed for the cointegration testing.

Table 13: Panel stationarity Test - First difference of Expenditure

Countries	Constant			Constant and trend		
	$k = 1$	$k = 2$	$k = 3$	$k = 1$	$k = 2$	$k = 3$
Czechia	0.195**	0.170	0.132	0.059**	0.094	0.093
Estonia	0.032	0.033	0.065	0.031	0.035	0.031
Hungary.	0.078	0.091	0.110	0.048*	0.089	0.107
Latvia	0.057	0.126	0.070	0.046	0.050	0.046
Lithuania	0.078	0.148	0.080	0.053*	0.097	0.060
Poland	0.075	0.044	0.055	0.045	0.038	0.051
Slovakia	0.122	0.136	0.137	0.047	0.083	0.088
Slovenia	0.064	0.131	0.097	0.060**	0.111*	0.096
Bulgaria	0.069	0.114	0.123	0.031	0.091	0.096
Romania	0.053	0.069	0.181	0.052*	0.040	0.068
Panel statistic	0.970 (0.166)	-0.827 (0.796)	-1.336 (0.909)	4.437*** (0.000)	1.674** (0.047)	1.034 (0.151)

Fourier Panel Stationarity Test for 10 CEEC under the Null hypothesis of Stationarity.
Critical values (obtained from Becker et al. 2006:289) for individual test statistics are as follows 0.1318 (10%), 0.1720 (5%), 0.2699 (1%) for $k=1$; 0.3150 (10%), 0.4152 (5%), 0.6671 (1%) for $k=2$; 0.3393 (10%), 0.4480 (5%), 0.7182 (1%) for $k=3$
Critical values for constant and trend are as follows 0.0471 (10%), 0.0546 (5%), 0.0716 (1%) for $k=1$; 0.1034 (10%), 0.1321 (5%), 0.2022 (1%) for $k=2$; 0.1141 (10%), 0.1423 (5%), 0.2103 (1%) for $k=3$

Significance at 10%, 5% and 1% are denoted by *, ** and *** respectively.

Table 14: Panel stationarity Test - First difference of Revenue

Countries	Constant			Constant and trend		
	$k = 1$	$k = 2$	$k = 3$	$k = 1$	$k = 2$	$k = 3$
Czechia	0.127	0.075	0.074	0.044	0.057	0.074
Estonia	0.093	0.082	0.117	0.051*	0.071	0.111
Hungary.	0.038	0.151	0.159	0.031	0.148**	0.153**
Latvia	0.063	0.062	0.107	0.047	0.052	0.056
Lithuania	0.081	0.122	0.079	0.061**	0.082	0.077
Poland	0.088	0.128	0.157	0.057**	0.056	0.060
Slovakia	0.108	0.531**	0.423*	0.076***	0.120*	0.099
Slovenia	0.061	0.108	0.117	0.057**	0.101	0.077
Bulgaria	0.180**	0.327	0.366*	0.027	0.149**	0.159**
Romania	0.038	0.065	0.083	0.035	0.048	0.041
Panel statistic	1.282 (0.100)	0.574 (0.283)	0.292 (0.385)	4.723*** (0.000)	2.965*** (0.002)	2.357*** (0.009)

Fourier Panel Stationarity Test for 10 CEEC under the Null hypothesis of Stationarity.
Critical values (obtained from Becker et al. 2006:289) for individual test statistics are as follows 0.1318 (10%), 0.1720 (5%), 0.2699 (1%) for $k=1$; 0.3150 (10%), 0.4152 (5%), 0.6671 (1%) for $k=2$; 0.3393 (10%), 0.4480 (5%), 0.7182 (1%) for $k=3$
Critical values for constant and trend are as follows 0.0471 (10%), 0.0546 (5%), 0.0716 (1%) for $k=1$; 0.1034 (10%), 0.1321 (5%), 0.2022 (1%) for $k=2$; 0.1141 (10%), 0.1423 (5%), 0.2103 (1%) for $k=3$

Significance at 10%, 5% and 1% are denoted by *, ** and *** respectively.

Table 15: Panel Stationarity test with sharp breaks - Robustness check

Panel A - Panel stationarity test (Revenue)		
Model	Level	First difference
Break (Homogeneous)	1.762(0.039)**	0.426(0.335)
Breaks (Heterogeneous)	2.975(0.001)***	0.152(0.440)

Panel B - Panel stationarity test (Expenditure)		
Model	Level	First difference
Break (Homogeneous)	1.687(0.046)**	0.137(0.446)
Breaks (Heterogeneous)	2.031(0.021)**	0.630(0.264)

Panel test by Carrion-i-Silvestre et al (2005). Reported test statistics and p-values in parenthesis. *,**,*** indicates rejection of the null hypothesis of "stationarity" at 10%. 5% and 1% respectively .

Table 16: Panel Cointegration test of cyclically adjusted revenue and cyclically adjusted spending (HP Filter used for detrending series)

Models	$Z_{\tau}(N)$		$Z_{\phi}(N)$	
	Value(τ)	P-Value	Value(ϕ)	P-value
No breaks	1.731	0.958	-2.409***	0.008
Level break	-1.984**	0.024	-2.490***	0.006
Regime shift	-6.050***	0.000	-7.182***	0.000
Number of observations	250		250	

Westerlund and Edgerton(2008) cointegration test. Maximum of three breaks are permitted. Displayed p-values is based on one-sided normal distribution test. *, ** and *** denotes rejection of the null hypothesis at 10%, 5% and 1% respectively. Maximum of 3 structural breaks in the cointegration relationship. Detrending of the series was done using the HP filter .

Table 17: Panel stationarity Test - Cyclically adjusted Expenditure

Countries	Constant			Constant and trend		
	$k = 1$	$k = 2$	$k = 3$	$k = 1$	$k = 2$	$k = 3$
Czechia	0.144*	0.165	0.132	0.065**	0.134**	0.126**
Estonia	0.080	0.401	0.385*	0.046	0.105*	0.092
Hungary.	0.157*	0.096	0.068	0.059**	0.100	0.126**
Latvia	0.106	0.202	0.155	0.045	0.080	0.073
Lithuania	0.089	0.172	0.146	0.072***	0.135**	0.126**
Poland	0.154*	0.466**	0.577**	0.047	0.074	0.073
Slovakia	0.095	0.258	0.206	0.063**	0.146**	0.126**
Slovenia	0.091	0.488**	0.332	0.070**	0.140**	0.093
Bulgaria	0.102	0.072	0.097	0.068**	0.072	0.065
Romania	0.064	0.087	0.110	0.049*	0.089	0.110
Panel statistic	2.491*** (0.006)	2.376*** (0.009)	1.461* (0.072)	7.017*** (0.000)	4.511*** (0.000)	2.885*** (0.002)

Fourier Panel Stationarity Test for 10 CEEC under the Null hypothesis of Stationarity.
 Critical values (obtained from Becker et al. 2006:289) for individual test statistics are as follows 0.1318 (10%), 0.1720 (5%), 0.2699 (1%) for $k=1$; 0.3150 (10%), 0.4152 (5%), 0.6671 (1%) for $k=2$; 0.3393 (10%), 0.4480 (5%), 0.7182 (1%) for $k=3$
 Critical values for constant and trend are as follows 0.0471 (10%), 0.0546 (5%), 0.0716 (1%) for $k=1$; 0.1034 (10%), 0.1321 (5%), 0.2022 (1%) for $k=2$; 0.1141 (10%), 0.1423 (5%), 0.2103 (1%) for $k=3$

Significance at 10%, 5% and 1% are denoted by *, ** and *** respectively.

Table 18: Panel stationarity Test - Cyclically adjusted revenue

Countries	Constant			Constant and trend		
	$k = 1$	$k = 2$	$k = 3$	$k = 1$	$k = 2$	$k = 3$
Czechia	0.160*	0.416**	0.519**	0.048*	0.070	0.060
Estonia	0.079	0.414*	0.249	0.044	0.042	0.089
Hungary.	0.040	0.398*	0.348*	0.039	0.058	0.103
Latvia	0.037	0.324*	0.356*	0.037	0.036	0.051
Lithuania	0.061	0.242	0.165	0.059**	0.145**	0.111
Poland	0.241**	0.434**	0.435*	0.078***	0.056	0.072
Slovakia	0.080	0.262	0.230	0.056**	0.149**	0.123*
Slovenia	0.097	0.237	0.261	0.075***	0.132*	0.066
Bulgaria	0.040	0.281	0.343*	0.037	0.035	0.071
Romania	0.083	0.246	0.338	0.054*	0.044	0.073
Panel statistic	1.524*** (0.064)	4.395*** (0.000)	3.768*** (0.000)	5.575*** (0.000)	1.985** (0.023)	1.682** (0.046)

Fourier Panel Stationarity Test for 10 CEEC under the Null hypothesis of Stationarity.
 Critical values (obtained from Becker et al. 2006:289) for individual test statistics are as follows 0.1318 (10%), 0.1720 (5%), 0.2699 (1%) for $k=1$; 0.3150 (10%), 0.4152 (5%), 0.6671 (1%) for $k=2$; 0.3393 (10%), 0.4480 (5%), 0.7182 (1%) for $k=3$
 Critical values for constant and trend are as follows 0.0471 (10%), 0.0546 (5%), 0.0716 (1%) for $k=1$; 0.1034 (10%), 0.1321 (5%), 0.2022 (1%) for $k=2$; 0.1141 (10%), 0.1423 (5%), 0.2103 (1%) for $k=3$

Significance at 10%, 5% and 1% are denoted by *, ** and *** respectively.

Table 19: Panel stationarity Test - First difference of cyclically adjusted Expenditure

Countries	Constant			Constant and trend		
	k = 1	k = 2	k = 3	k = 1	k = 2	k = 3
Czechia	0.224**	0.255	0.298	0.140***	0.142**	0.137*
Estonia	0.041	0.105	0.061	0.037	0.043	0.045
Hungary.	0.059	0.141	0.144	0.045	0.141**	0.131*
Latvia	0.076	0.075	0.068	0.056**	0.063	0.068
Lithuania	0.121	0.211	0.194	0.119***	0.126*	0.118*
Poland	0.193*	0.191	0.198	0.051*	0.120*	0.144**
Slovakia	0.044	0.304	0.204	0.036	0.036	0.063
Slovenia	0.138*	0.189	0.147	0.108***	0.137**	0.074
Bulgaria	0.102	0.073	0.142	0.049*	0.051	0.061
Romania	0.042	0.147	0.055	0.038	0.048	0.044
Panel statistic	2.237** (0.013)	0.668 (0.252)	-0.086 (0.532)	9.258*** (0.000)	3.130*** (0.001)	2.196** (0.014)

Fourier Panel Stationarity Test for 10 CEEC under the Null hypothesis of Stationarity.
 Critical values (obtained from Becker et al. 2006:289) for individual test statistics are as follows 0.1318 (10%), 0.1720 (5%), 0.2699 (1%) for k=1; 0.3150 (10%), 0.4152 (5%), 0.6671 (1%) for k=2; 0.3393 (10%), 0.4480 (5%), 0.7182 (1%) for k=3
 Critical values for constant and trend are as follows 0.0471 (10%), 0.0546 (5%), 0.0716 (1%) for k=1; 0.1034 (10%), 0.1321 (5%), 0.2022 (1%) for k=2; 0.1141 (10%), 0.1423 (5%), 0.2103 (1%) for k=3

Significance at 10%, 5% and 1% are denoted by *, ** and *** respectively.

Table 20: Panel stationarity Test - First difference of cyclically adjusted revenue

Countries	Constant			Constant and trend		
	k = 1	k = 2	k = 3	k = 1	k = 2	k = 3
Czechia	0.060	0.092	0.168	0.058**	0.068	0.067
Estonia	0.043	0.073	0.097	0.039	0.067	0.086
Hungary.	0.042	0.071	0.121	0.040	0.071	0.122*
Latvia	0.093	0.096	0.092	0.093***	0.096	0.088
Lithuania	0.166*	0.496**	0.303	0.103***	0.130*	0.104
Poland	0.217**	0.187	0.286	0.033	0.071	0.089
Slovakia	0.102	0.244	0.214	0.102***	0.104*	0.095
Slovenia	0.131	0.172	0.185	0.114***	0.197**	0.088
Bulgaria	0.1068	0.085	0.117	0.063**	0.077	0.116*
Romania	0.120	0.071	0.052	0.039	0.047	0.042
Panel statistic	2.244** (0.012)	0.421 (0.337)	0.190 (0.425)	9.430*** (0.000)	3.306*** (0.000)	2.275** (0.011)

Fourier Panel Stationarity Test for 10 CEEC under the Null hypothesis of Stationarity.
 Critical values (obtained from Becker et al. 2006:289) for individual test statistics are as follows 0.1318 (10%), 0.1720 (5%), 0.2699 (1%) for k=1; 0.3150 (10%), 0.4152 (5%), 0.6671 (1%) for k=2; 0.3393 (10%), 0.4480 (5%), 0.7182 (1%) for k=3
 Critical values for constant and trend are as follows 0.0471 (10%), 0.0546 (5%), 0.0716 (1%) for k=1; 0.1034 (10%), 0.1321 (5%), 0.2022 (1%) for k=2; 0.1141 (10%), 0.1423 (5%), 0.2103 (1%) for k=3

Significance at 10%, 5% and 1% are denoted by *, ** and *** respectively.

Figure 5: Structural breaks in government Expenditure to GDP ratio - red vertical lines indicate breakpoints

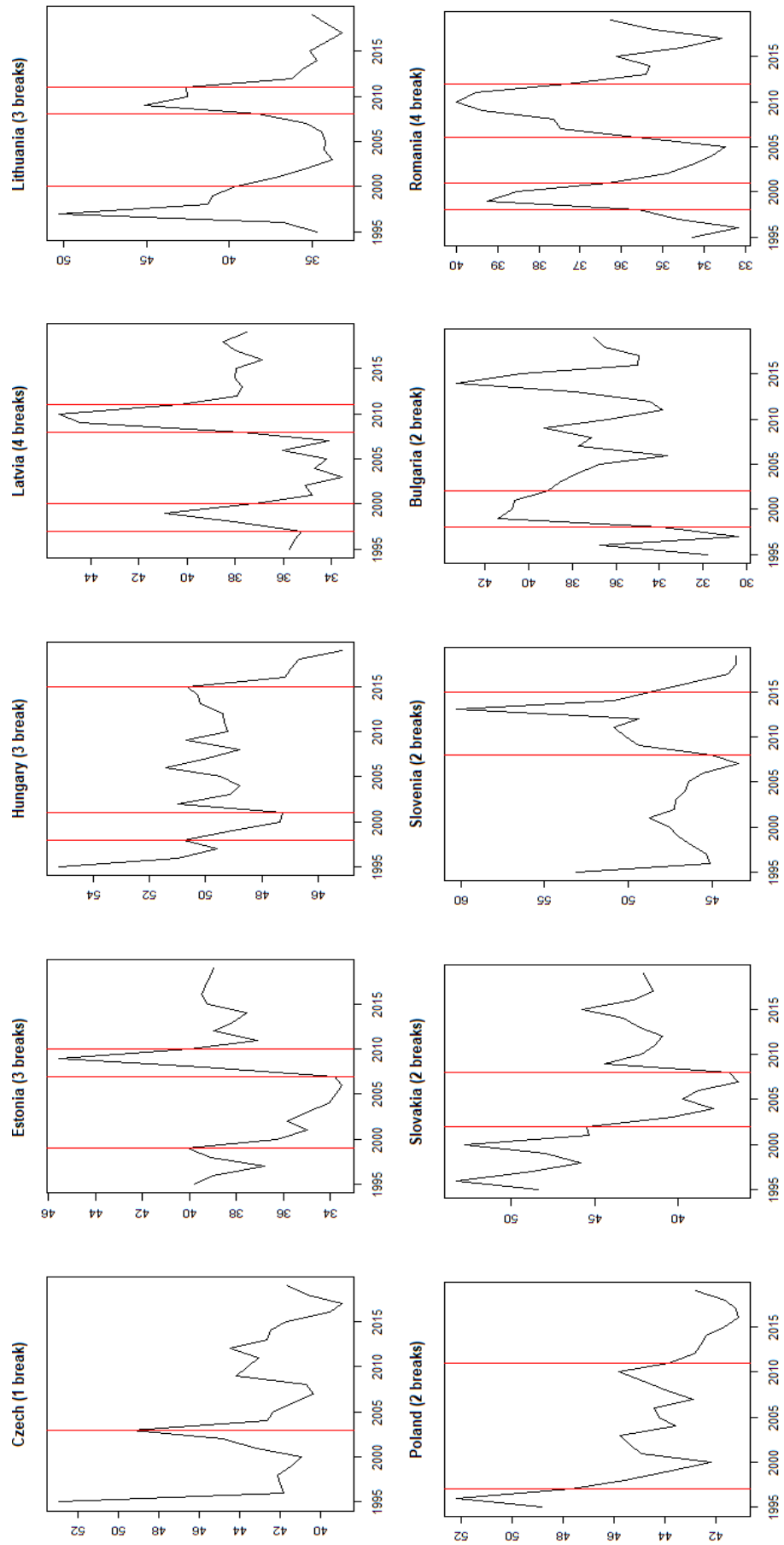


Figure 6: Structural breaks in government Revenue to GDP ratio - red vertical lines indicate breakpoints

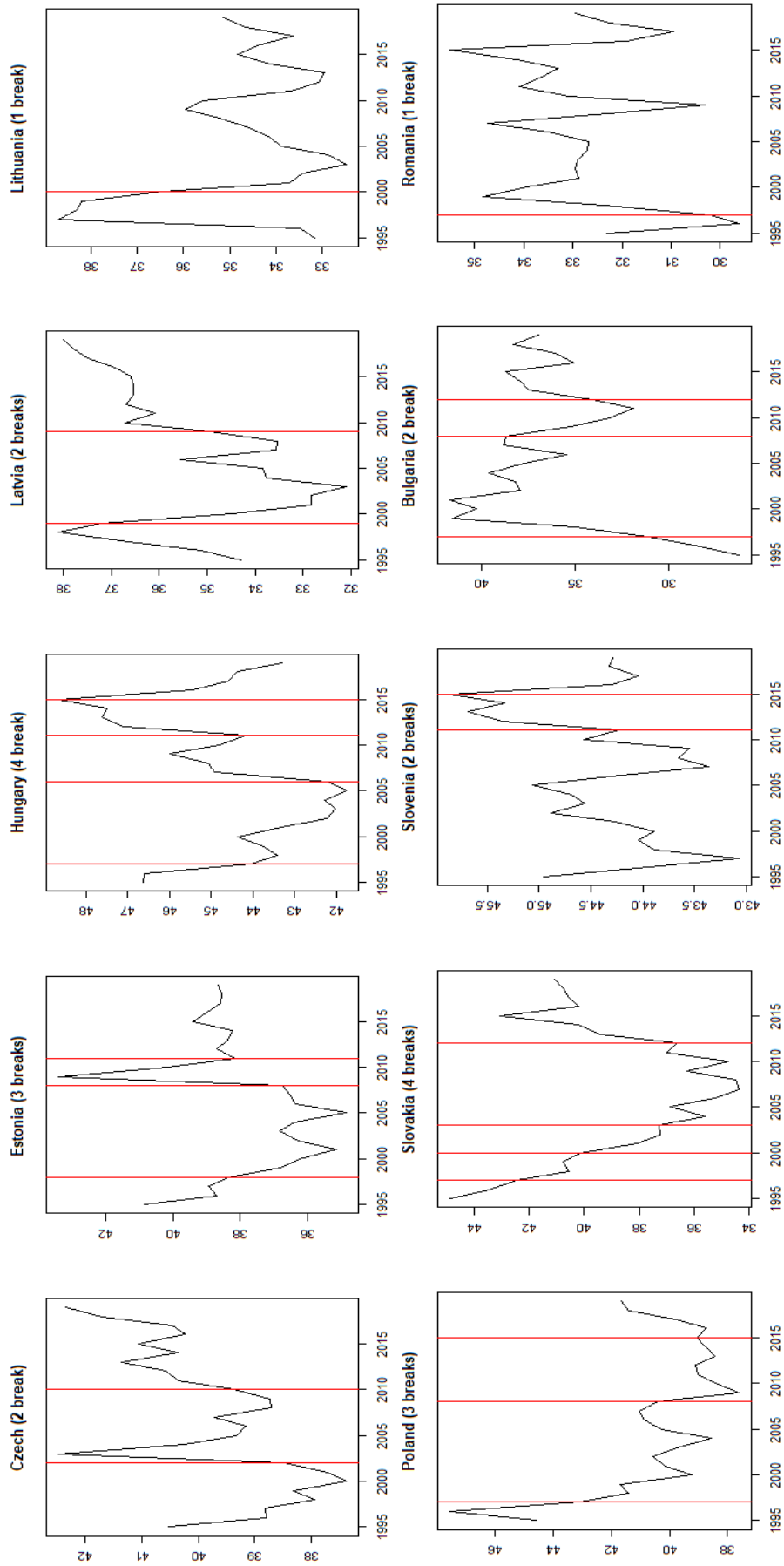


Figure 7: Hamilton filter estimates of the trend component (black) versus actual series (red) - Expenditure to GDP ratio

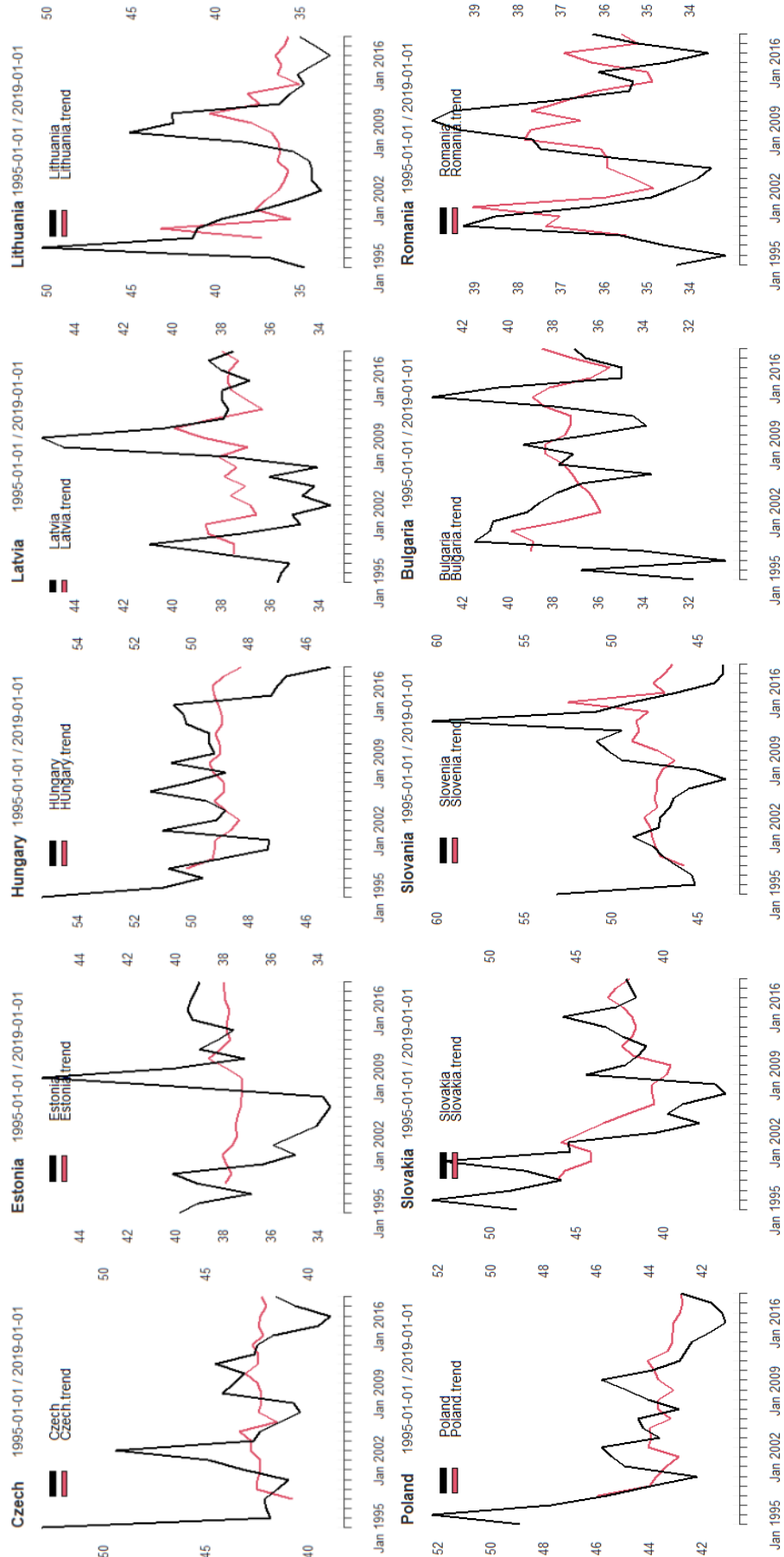


Figure 8: Hamilton filter estimates of the trend component (red) versus actual series (black) - Revenue to GDP ratio

