Does the structural budget balance guide fiscal policy pro-cyclically? Evidence from the Finnish Great Depression of the 1990s

Kuusi, Tero

The Research Institute of the Finnish Economy

26 February 2018

This is a pre-copyedited, preprint version of an article that is published as Kuusi, Tero (2017), Does the structural budget balance guide fiscal policy pro-cyclically? Evidence from the Finnish Great Depression of the 1990s, National Institute Economic Review, No 239 February 2017. The published version can be found here:

http://journals.sagepub.com/doi/pdf/10.1177/002795011723900111

This research has been supported by European Union's Horizon 2020 research and innovation programme under grant agreement No 64926, project FIRSTRUN (Fiscal Rules and Strategies under Externalities and Uncertainties).

Does the structural budget balance guide fiscal policy pro-cyclically? Evidence from the Finnish Great Depression of the 1990s

Tero Kuusi[1]

Abstract:
In this article, I evaluate the challenges related to the European Commission's output gap method of calculating the structural budgetary position, and assess its bottom-up alternatives in the EU's fiscal framework using the Finnish data for the years 1984-2014. The results reinforce the impression of the limited capacity of the output gap method to predict cyclical changes in real time and suggest that using the output gap method to steer fiscal policy tends to lead to a procyclical policy (stimulus in upturns and austerity in downturns). The bottom-up assessment methods that are based on discretionary fiscal policy measures appear to work better, and using them to steer the fiscal policy could make the policy more countercyclical.

Keywords: structural budget balance, output gap, fiscal stance, discretionary fiscal effort
JEL Classifications: E62, H60

1. Introduction

[1] This study has been funded from appropriations for assessment and research activities in support of the Government's decision-making, and the Horizon 2020 Framework Programme of the European union under the grant agreement number 649261 (the Firstrun project). I would like to thank the following persons for their expert comments: Zsolt Darvas, Marketta Henriksson, Hannu Kaseva, Markku Kotilainen, Harri Kähkönen, Tuomo Mäki, Niku Määttänen, Veli-Arvo Tamminen,Jenni Pääkkönen,Tarmo Valkonen, Vesa Vihriälä and the seminar audience during internal seminars within ETLA, the Finnish Ministry of Finance, as well as the FIRSTRUN workshop and the joint EPC – ECFIN – JRC Workshop "Assessment of the real time reliability of different output gap calculation methods" in 2015. The author assumes sole responsibility for the contents of this paper and any errors therein.

The structural budget balance (SB) measures the budgetary position of public finances, when the effects of economic cycles and one-off expense and income items are eliminated. It has received a central role in the EU's fiscal policy legislation framework. In the corrective arm of the Stability and Growth Pact, SB will help steer the removal of excessive deficit. In the preventive arm of the Stability and Growth Pact, it specifies the government's general medium-term budgetary objective. In principle, the use of SB clarifies the execution of fiscal policy and its control. Public finances should react to shocks of a cyclical nature with automatic stability measures, and in principle, such measures should be allowed to work in spite of the short-term costs inflicted on public finances. However, if the SB worsens, the related change in fiscal policy can be interpreted as independent of economic cycles, and should be corrected at least in cases where the sustainability of the public finances is in danger. Without steering produced by fiscal policy indicators such as SB, uncertainty about the nature of shocks can easily lead to contradictory policy recommendations, which could, in the worst case, paralyse fiscal policy.[2]

Despite its conceptual clarity, it is challenging to measure SB in practice. This is especially true of the so called output gap-based methods of measuring SB. The methods require assessments on several quantities that are difficult to measure (see for example Mourre et al. 2013; Havik et al. 2014). First, the output gap must be defined, i.e. the difference between actual economic activity and potential economic activity must be estimated. The structural budgetary balance is calculated next, taking account of the historical sensitivity of tax revenue and public expenditure to fluctuations in the output gap. The resulting assessments of the effects of fiscal cycles on the budgetary position of public finances in different countries have been criticised as inadequate during the recent financial and debt crisis. If this is the case, fiscal policy reliant on such indicators is in danger of becoming procyclical.[3]

In this article I assess the challenges in the European Commission's method of calculating SB and consider methodological alternatives to the output gap method. The perspective adopted is that of recent Finnish economic developments between the years 1984-2014. The time period provides rapid swings in Finland's business cycles, from fiscal overheating in the late 1980s through deep crisis in the early 1990s to recovery and growth since the mid-1990s, the IT bubble in the early 2000s, and finally the Great Recession in the late 2000s. The well-documented time period makes

[2] While this paper focuses on the use of numerical rules to guide countercyclical policy-making (see, e.g, Portes and Wren-Lewis, 2015; Bergman and Hutchison, 2015, Sacchi and Salotti, 2015), that is, of course, only one factor that motivates their existence. A large literature on the governance of fiscal policy stresses the role of fiscal rules in curtailing political incentives to adopt policies likely to benefit the policy-makers rather than the interests of the economy (Begg, 2016). It encompasses issues such as the nature of the 'contract' between citizens as principals and their governments as their agents, the most appropriate design of institutions, and transparency (Besley 2007; Hallerberg et al. 2007; Begg 2014). The recent research finds evidence that sustainable public finances in Europe may be associated with strong fiscal rules, and that fiscal rules and government efficiency may be institutional substitutes in terms of promoting fiscal sustainability (Bergman et al. 2016).

[3] For example, Lane et al. (2012) argue that prior to the Eurozone crisis, financial policy was excessively based on output gap estimates, without taking into account the risks associated with external imbalances, credit expansion, debt overhang in various sectors and housing price trends. On the other hand, after the crisis broke out, concerns were expressed that the output gap-based assessment of the correction needed for SB had not produced the correct picture of adjustments made in the public finances (European Commission, 2013B).

it possible to examine in great detail both the functioning of European Commission's method of calculating SB, as well as alternative indicators that might serve as inputs for tuning fiscal policy.

With regard to the output gap method, I calculate historical estimates for two key components of the output gap (structural unemployment and the potential level of total factor productivity) in 1984–2014. I also examine the plausibility of the Commission's current estimates by comparing them to observations in earlier literature. In addition, I evaluate the method at various points in real time – that is without information on the future development of economy that would be available later.

The results reinforce the impression of the limited capacity of the output gap method to predict cyclical changes in real time. Therefore, its use for steering fiscal policy - in the EU's fiscal framework for instance – tends to lead to a procyclical fiscal policy (stimulus in upturns and austerity in downturns). For example, fiscal policy guided by an output gap-based SB does not react in a contractionary manner during the economic upswing in the 1980s and early 2000s. On the contrary, the indicator permits fiscal policy to be more expansive than the actual policy, if it had been calculated with real time data. Besides, an output gap-based SB indicator ignores the fairly strong contractionary measures in fiscal policy implemented in the crisis of the early 1990s, thereby leading to even more contractionary policy.

Based on the results, it appears that the method currently used by the Commission may also be hypersensitive to changes in economic trends due to methodological reasons. In particular, estimates about structural unemployment in the recession of the 1990s that have increased to a quite high level indicate that the indicator could overreact to economic cycles. One explanation for the behaviour is provided by the statistically problematic constraints on the parameter estimates that are imposed when applying the estimation.

As methodological alternatives to the output gap method, I will review other fiscal policy evaluation methods used in the EU's legislation framework: the expenditure benchmark in the preventive part, and a bottom up assessment method in the corrective part of the Stability and Growth Pact. It is important to review alternative methods, since they measure the budgetary position using fairly different criteria. Unlike the SB, both the expenditure rule and the bottom up assessment evaluate potential production in the medium term. Cyclical expenditure items are subtracted from public expenditure more directly than in assessments based on an output gap and standard cyclical elasticity, and the revenue trend is measured based on the observed decisions on a revenue basis and assessments of their effects.

In practice, alternative indicators already form part of the EU's control of fiscal policy. An understanding of the practicality of the various methods is also necessary due to the fact that the EU rules on fiscal policy leave much room for selecting the indicator used to guide fiscal policy (although the output gap method still plays a fairly central role within the rules). In the preventive arm of the Stability and Growth Pact, the actualisation of the medium-term budgetary objective is assessed not only by output gap-based SB, but also by the expenditure rule. According to the

expenditure rule, public expenditure may only grow at the same rate as the potential medium-term GDP used as the reference. In the excessive deficit procedure of the corrective arm, the effectiveness of corrective measures is assessed not only via the SB, but also in terms of the number of discretionary measures in question. In practice, such an assessment is based on a method that resembles the expenditure rule very closely. Using this method, cyclical items are eliminated from the expenditure trend, which is then compared to the medium-term growth of potential production, taking account of changes in the revenue basis (bottom up assessment).

For the analysis of alternative methods, I have collected a new historical time series on the effects of the changes on the revenue basis of the entire public economy (the state, local administration and social funds). Using the data, I analyse how the alternative methods would have worked over the last three decades.

The results are encouraging. Using either the expenditure rule or the bottom up assessment to guide fiscal policy results in a more countercyclical policy than the output gap-based SB. Fiscal policy based on the expenditure rule is contractionary, especially during the lead up to the 1990s crisis, which could have helped to alleviate the crisis and increase the margin for recovery measures while it was taking place. On the other hand, based on a discretionary bottom up assessment, the contractionary fiscal policy practised from 1992 onwards is sufficient, and unlike the output gap-based SB, the method does not generate additional contractionary pressures. It is also noteworthy that in spite of their different assumptions, the methods provide a fairly uniform view of the magnitude of discretionary measures.

In Section 2 of this article, I present the used methods. In Section 3, the applied data is introduced, while in Section 4 the results of the analysis are reported. Section 5 concludes.

2. Methodology

In this section, I briefly present the output gap-based structural balance and its discretionary alternatives within the EU's fiscal policy legislation framework.

2.1. Structural balance with the Commission's output gap method

In the European Commission's calculation method the structural balance (SB) is calculated on the basis of estimates about the historical sensitivity of tax revenue and public expenditure to fluctuations in the output gap. This is assessed as the difference between the actual fiscal position and the cyclic effects as relative to the GDP:

$$SB_t = \frac{R_t - G_t}{Y_t} - \epsilon * OG_t - OO_t,$$

where R_t is public sector revenue, G_t is public sector expenditure and Y_t is the nominal GDP at year t. The cyclic correction is the product of the output gap (OG_t) and the elasticity between the output gap and budgetary balance ϵ. In the method used by the Commission, the output gap is determined in proportion to the production potential of the entire national economy, and semi-elasticity ϵ is assumed to be a constant. In addition, the budgetary balance is adjusted in proportion to GDP by using the effect of certain one-off revenue and expenditure items (OO_t). Mourre et al. (2013) reviews the semi-elasticity ϵ calculation method in more detail.

Currently, most international institutions (OECD, IMF, European Commission) calculate potential output using the production function method, which enables the efficient utilisation of the available research information on production technology and the behaviour of various factors of production during the assessment of the cyclic phase of the economy. The idea is to aggregate a comprehensive view of the production capacity of the economy (potential production function), based on an economic theory and observations of the state of the various components.

In the method applied by the European Commission (see Havik et al. 2014), the production function is assumed to follow the Cobb-Douglas form and it can be presented as

$$Y_t = (U_{Lt}L_tE_{Lt})^{\alpha}(U_{Kt}K_tE_{Kt})^{1-\alpha} = TFP_t{K_t}^{\alpha}{L_t}^{1-\alpha},$$

where Y_t is total production, L_t total labour input, and K_t physical capital stock. The use of each production factor is controlled by their utilisation rate (U_{Lt}, U_{Kt}) and the efficiency of use (E_{Lt}, E_{Kt}). The parameter α measures the expenditure share of labour input of all inputs. Labour input is measured as the total number of work hours, and capital is measured as the amount of capital services, divided into buildings and equipment. The Cobb-Douglas production function allows total factor productivity to be examined separately as the weighted product of efficiency and the utilisation rate.

$$TFP_t = (U_{Lt}E_{Lt})^{\alpha}(U_{Kt}E_{Kt})^{1-\alpha},$$

The output gap can be divided into different components. When the potential magnitude of the components of the production function is known, the percentual deviation from potential can be approximately estimated as the difference between the logarithms of the components

$$OG_t = LN(Y_t) - LN(Y_t^{pot}) = LN(TFP_t) - LN(TFP_t^{pot}) + (1-\alpha)(LN(L_t) - LN(L_t^{pot})).$$

It is worth noting that, in the output gap calculation, the capital stock is not adjusted separately in line with the phase of the economic cycle. Moreover, the quantity of the potential workforce is divided further into several components. This corresponds to the potential workforce adjusted

based on the level of structural unemployment, $NAWRU_t$. The potential workforce is the product of the size of the population of working age people POP_t^W, the average level of participation $PART_t^{pot}$ and working hours per employee H_t^{pot}.

$$L_t^{pot} = POP_t^W PART_t^{pot}(1 - NAWRU_t)H_t^{pot}.$$

The cyclical adjustment of participation and working hours is based on a statistical HP filter. Thus, the assessment of trends does not include a separate economic theory. The population of working age is measured based on the actual number of people of working age.

Here, the focus is particularly on the methods of estimating structural unemployment and total factor productivity.[4] With regard to distinguishing the cyclical and structural components of unemployment, the Commission uses a general labour market framework whose features are ultimately estimated based on the data and correspond to the predictions of various labour market theories (see Havik et al. 2014). Outside the long-term equilibrium, the short-term state of the labour market can be assessed using the Phillips curve. This curve describes the inverse relationship between inflation and cyclical unemployment. Key factors affecting the curve include assumptions about the creation of expectations. The total factor productivity term is also broken down into a cyclical and structural component, but unlike for unemployment, no precisely described theoretical model can be invoked to justify the breakdown. Instead, it is assumed that the cyclical term depends on the underutilisation of economic resources, which is measured using the capacity utilisation rate series and by making assumptions about the duration of the effects of the underlying shocks. [5]

2.2.Critique of the SB

The measurement of output gap-based structural balance has been studied quite extensively in the literature, and an increasing number of reservations have been raised concerning its use. The estimation of the output gap is highly sensitive to changes in estimates over time, both due to genuine uncertainty and to the difficulty of selecting the right model (e.g. Orphanides and van Norden, 2002; Rünstler, 2002; Planas and Rossi, 2004; Golinelli, 2008; Marcellino and Musso, 2010; Bouis et al., 2012).

[4] The components play a central role in the output gap method and offer the greatest opportunities for a review from an economics point of view.

[5] Conclusions about unobservable structural changes in these components are made using the maximum likelihood method, a Bayesian method of calculation, and the Kalman filter. A more detailed description of the method is presented by Kuusi (2015), Planas and Rossi (2004), Planas and Rossi (2014) and Havik et al. (2014)

The uncertainties relates to various components of the output gap.[6] First, it concerns the form of the production function. For example, in the Finnish case Jalava et al. (2006) state that the Cobb-Douglas production function (whereby the nominal shares of the factors remain constant) may be statistically applicable in the long-term, but not completely adequate for describing Finland's production since World War II. Luoma and Luoto (2010) are of the opinion that a more suitable production function would be the CES (Constant Elasticity of Substitution) production function, whereby the nominal proportions of production inputs may vary as their relative prices change.[7] Nevertheless, when applied during a crisis, the Cobb Douglas production function can be argued to provide a good estimate of the CES production function, even if the predictions generated by the Cobb Douglas production function would not work in the long term. (Havik et al., 2014). In any case, in the short run it is very difficult to assess technological development supporting various production factors and the change in respective input proportions. The effect of trends that are often weak but that affect the production function in the long term is dominated by the effect of a crisis on the profitability, efficiency and product demand of companies.

The estimation of the individual components of the output gap also involves uncertainties. For example, with respect to the measurement of cyclical unemployment, the recent literature suggests that the behavior of inflation does not necessarily correspond to the New-Keynesian Phillips curve during major crises, even if it includes backward-looking elements, such as a lagged inflation term. For example, Stock and Watson (2010) are of the opinion that, in the US, an increase in unemployment does decrease inflation, but this effect wears off when a higher level of unemployment has lasted for 11 quarters. One of the underlying causes of this could be anchored inflation expectations, whose effects during the euro crisis are a topic of discussion, see for example Krugman (2013). Wage frictions (for example, pressure not to reduce nominal wages) can affect the relation between inflation and unemployment in such a way that it does not correspond to the New-Keynesian Phillips curve. (Daly and Hobijn, 2013). In the Finnish case, there is clear evidence of fairly substantial wage inelasticity in the crisis of the early 1990s (Gorodnichenko et al., 2012).

Another key challenge is the estimation of the total-factor productivity (TFP) gap. The interpretation and forecasting of TFP growth can be problematic as it is measured as a residual growth of output after the influence of production factor growth is accounted for. Thus, TFP growth may result from multitude of factors, such as capacity utilization, increasing returns to scale, mark-ups due to imperfect competition, or gains from sectoral reallocations, as well as

[6] While traditionally the output gap estimation has been based on the trend estimation, here the focus is on the production function based estimation of the output gap that most institutions currently use (OECD, IMF, European Commission). Murray (2014) reviews various trend estimation methods.

[7] A key question when selecting a production function is that of how technological development affecting production factors – capital and labour – changes the quantity of inputs adjusted for technological development, and their nominal proportion in production. Research on Finland indicates that, in the long-term, the proportions of the inputs change: the input proportion of the production factor that becomes cheaper (capital) reduces in proportion to the factor that becomes relatively more expensive, and on the other hand, technological development may support the growth of the amount of capital more than it raises efficiency in the utilisation of labour.

measurement errors of the inputs.[8] Furthermore, TFP is subjected to major trend volatility that makes it difficult to assess its potential level. In this respect, Finland provides an illustrative example. In each major economic crisis of the last 30 years, the Finnish economy has suffered from structural surprise shocks that have persistently affected productivity (the Soviet trade collapse in the early 1990s and the recent collapse of the Nokia in the late 2000s). The shocks have been largely unanticipated and their aggregate productivity impacts have been hard to predict. The recent crisis shows that the uncertainty regarding the long-term productivity growth is not unique to Finland. [9]

All in all, a look on the Commission's method in the present crisis confirms that revisions of the output gap have been large. For example, Virkola (2014) reviews the revisions made to the European Commission's output gap methods, and reports that the changes to output gap estimates in 2000–2013 amounted to 1.5 percentage points on average during the crisis.

Challenges associated with the calculation of the output gap-based SB are not, however, limited to the difficulty of measuring the output gap, but also relate to the difficulty of modelling the reactions of the public economy to cyclic shocks. Firstly, a cycle-independent budget should not contain individual expenditure and revenue items that have no clear connection to the long-term balance. Although it is easy to eliminate one-off items from the budget in principle, problems occur when trying to define which items are temporary or large enough (European Commission 2006). Secondly, the budget balance of the public finances can depend on fluctuations in asset and commodity prices that correlate only weakly with economic cycles (see for example Eschenbach and Schuknecht 2002, Price and Dang 2011). In addition, economic crises and their aftermaths are associated with structural and legal reforms that do not treat every sector and public finance revenue base equally. Taking them into account requires an alternative approach to SB calculation, since calculations based on an aggregated output gap assume that economic upswings and downswings are symmetrical and thus neutral towards sources of tax revenue (Kremer et al. 2006; Morris 2007; Wolswijk 2007; Barrios and Fargnoli 2010).

2.3.The alternative indicators

I evaluate alternative indicators that have recently been presented as solutions to the problems presented above within the EU's fiscal policy legislation framework. These comprise the expenditure rule within the preventive arm of the SGP, which is defined in the Commission's vade mecum guidelines (2013A, 2016). The purpose of the expenditure rule is to ensure that the

[8] In this respect, one of the problem is the availability of good-quality data on capital that has historically been limited. (see, e.g, Bryson and Forth (2015) for the UK). While this paper abstracts from the problems regarding the measurement of the (productive) capital stock, a good reference is D'Auria et al. (2010)

[9] For example, before and during the Great Recession, the US real-time data obscured the slowdown in trend, and overstated productivity's strength early in the recession. Almost every revision since 2005 has lowered the path of labor productivity (Fernald 2014). Similar patterns are widely seen in other countries (UK, Bryson and Forth 2015; Europe, Summers, 2014)

countries remain committed to their medium-term objectives (MTO) or a path of adjustments leading to it. On the other hand, the excessive deficit procedure in the SGP's corrective arm assesses the outcomes of actions that seek to correct the budgetary position by means of a bottom up assessment. It resembles very closely the expenditure rule in the preventive arm in methodological terms. The latter indicator is discussed by the European Commission (2013B) and Carnot and de Castro (2015), among others.

The starting point in both alternative indicators is the direct analysis of detected policy changes instead of indirect assessments based on the output gap method. In principle, it is easy to monitor changes in economic policy on the revenue side: economic policy is essentially neutral if no new decisions are made. The combined effects of new decisions can be interpreted as a change in fiscal policy.

On the other hand, there is no corresponding distinct neutral reference point on the expenditure side, as changes in the expenditures involve more automatic responses to the economic conditions, but the growth in expenditure must somehow be quantified in reference to other development in the aggregate economy. Changes in fiscal policy are measured based on the growth rate of aggregated expenditures, with various cyclical items being eliminate, as relative to the potential medium-term growth in GDP[10]. A fiscal policy can be interpreted as neutral if it will not change the expenditure proportion of GDP according to the adjusted expenditure in the medium term. On the other hand, if the adjusted expenditure growth rate exceeds the potential growth of GDP in the medium term, the fiscal policy must be interpreted to have changed, particularly if the difference will not be compensated with discretionary measures on the revenue side.

In the following, I will examine alternative indicators in more detail. In the case of the expenditure rule, revenue base changes and various cyclical items are subtracted from public expenditure

$$E_t = G_t - INT_t - EU_t - (I_t - I_t^{AVE}) - UC_t$$

[10] However, it must be noted that the Commission's method of measuring potential production is also applied when making these longer-term assessments. This could still present a problem, especially since the output gap method includes an assumption on the closing of the output gap, which could also generate biased forecasts in the medium term (Timmermann 2006). An alternative method of measuring potential production could, for example, lie in the long-term growth forecasting method used by the US Congressional Budget Office (CBO) (Schackleton 2013; Hetemäki 2015). In the case of Finland, on the other hand, shocks have often occurred at the sector level. Thus, it may be sensible to consider an alternative whereby the development of production is estimated from the sector level upwards, using growth accounting or sector-level growth models (Pohjola 2011; Kuusi 2013; Fernald 2014).

where in year t, G_t is total public expenditure, INT_t interest expenses, EU_t the country's share of EU structural fund projects, I_t public investment expenditure, I_t^{AVE} average public investment expenditure in the current and three previous years, and UC_t cycle-related variation in unemployment expenditure. Unemployment expenditures due to economic cycles are assessed based on an estimate of the magnitude of cyclical unemployment (derived from the magnitude of structural unemployment) and average unemployment expenses per unemployed person.

The change in adjusted aggregated expenditures is calculated further, taking account of the discretionary change in revenue N_t^R (and certain expenses funded by earmarked revenue) in such a way that the proportional change in expenses is

$$\frac{\Delta E_t}{E_{t-1}} = \frac{E_t - N_t^R - E_{t-1}}{E_{t-1}}$$

The growth rate of expenses is deflated using the price change in GDP. Using the method of calculating the expense rule, inflation is measured as the average of the Commission's previous year's spring and autumn inflation forecasts for the current year. Let us express the real change as $\frac{\Delta e_t}{e_{t-1}}$.

The estimate of growth potential is based on the potential change in the level of production by the aggregate economy in the medium term. When the growth rate of expenditure equals the potential growth rate of production, the economy does not include a tendency to increase or decrease public demand in proportion to GDP in the medium term. Based on the Commission's suggestion, the potential growth rate is defined as the average based on observations of the growth rate of potential GDP during the last five years and forecasts of the growth rate for four years into the future:

$$\frac{\Delta_t^{pot} e_t}{e_{t-1}} = \left(\left(\frac{Y_{t+4}^*}{Y_{t-5}^*}\right) - 1\right)^{\frac{1}{10}},$$

where Y_t^* is potential (real) production at a particular point of time t.

When the adjusted expenditure aggregate has been calculated, its real growth $\frac{\Delta e_t}{e_{t-1}}$ can be compared to the growth potential of the aggregate economy $\frac{\Delta_t^{pot} e_t}{e_{t-1}}$.. A useful result is that the growth of expenditure aggregate must undershoot the reference growth rate by $x * \frac{1}{E_t/Y_t}$, to have the corresponding proportion of expenditure to GDP fall by x per cent, where E_t/Y_t is the nominal GDP proportion of the expenditure variable used.

In a bottom up estimate, the definition of the adjusted expenditure aggregate is slightly different to the expenditure benchmark. The expenditure aggregate is defined by first subtracting the non-

discretionary unemployment expenditure (G_t) interest expenses of public bodies (U_t^{nd}) and one-off expenditure items (I_t) from the total expenditure of public bodies (OO_t):

$$E_t^{BU} = G_t - U_t^{nd} - I_t - OO_t.$$

The change rate of expenditure is estimated as above

$$\frac{\Delta E_t^{BU}}{E_{t-1}^{BU}} = \frac{E_t^{BU} - N_t^R - E_{t-1}^{BU}}{E_{t-1}^{BU}}.$$

The *discretionary fiscal effort* (DFE_t) resulting from the nominal difference between the expenditure variable and reference growth indicates their impact on the change in the proportion of expenses in GDP between years t and t-1. I define DFE in the same way as the European Commission (2013B) and Carnot and de Castro (2015), as the difference between growth rates divided by the GDP ratio of the expenditure aggregate, as follows:

$$DFE_t = -\frac{\frac{\Delta E_t^{BU}}{E_t^{BU}} - \frac{\Delta_t^{pot} E}{E_{t-1}}}{\frac{Y_t}{E_t^{BU}}} = -\frac{E_t^{BU} - N_t^R - E_{t-1}^{BU}}{Y_t} + \frac{\frac{\Delta_t^{pot} E}{E_{t-1}} E_{t-1}^{BU}}{Y_t}$$

$$= \frac{N_t^R}{Y_t} - \frac{E_t^{BU} - E_{t-1}^{BU} - \frac{\Delta_t^{pot} E}{E_{t-1}} E_{t-1}^{BU}}{Y_t} = DFE_t^R + DFE_t^E,$$

where the reference growth of potential production is now defined as nominal $\frac{\Delta_t^{pot} E}{E_{t-1}} = \left(1 + \frac{\Delta_t^{pot} e_t}{e_{t-1}}\right) * \frac{P_t}{P_{t-1}} - 1$.. In the last breakdown, the indicator is further divided into the impact of revenue base changes (DFE_t^R) and the change in expenditure related to potential (DFE_t^E).

Subject to reservations due to the differences in the methods, both the DFE indicator and SB can measure the same cycle-independent change in the budgetary position. If the DFE indicator is positive by 1 percentage point, the growth rate of expenditure (with an adjusted expense aggregate and taking the revenue side into account), is estimated to be so slow that the budgetary position is strengthened on a discretionary basis by 1 percentage point.

The theoretical connection between the output gap-based SB and the DFE indicator defined by aggregated expenditures used in a bottom up assessment has been reviewed by the European Commission (2013B, box III.2.1) and Carnot and de Castro (2015, Appendix 1). In principle, the indicators are equivalent: During long-term growth equilibrium, where the elasticity of revenue and expenditure items are close to the averages estimated using the fixed elasticity method and economic growth remains stable, very similar results should be yielded by the different methods. However, differences may appear in the case of a large shock. Based on the breakdowns of the

two indicators, it becomes apparent that the differences on the revenue side are explained by changes in expenditure elasticity in cycles (such as windfall revenue), deviations in income class proportions from their fixed shares according to the fixed elasticity method, and changes generated by potential output in the long-term ratio of revenue and GDP. Of the above, changes in cyclical elasticities associated with windfall revenue are by far the most significant explanatory factor according to Carnot and de Castro (2015). On the expenditure side, the differences are mainly explained by unemployment expenditure that cannot be directly attributed to cycles, differences in methods of measuring potential output, or interest expenses.

3. Data

3.1. Data used in the evaluations of the output-gap method

I mainly use the data from the European Commission´s autumn 2014 forecast as material. This comprises a time series on unemployment ranging from 1963 to 2016. The data between 2014 and 2016 comprises forecasts. The inflation variable used in the Phillips curve is the change in unit labour costs. The unit labour cost is equal to wage inflation less the labour productivity growth rate and the change in consumer prices. The material extends up to 2014, while the data for 2014 is the forecast by the Commission. [11]

In addition, the data consists of the total-factor productivity series which the Commission calculates using real GDP, capital and labour series, as well as their expenditure shares of the inputs in production. The figures for 2014 to 2016 are based on the Commission's forecasts.

It also includes a capacity utilisation rate series, which is a collection of business cycle indicators describing economic activity (Havik et al., 2014). The series' components consist of the industrial capacity utilisation rate as well as service sector and construction sector confidence indicators. The indicators are weighted with the shares of total output of the economy attributable to different sectors, and their standard deviations are normalised in such a way that the deviations correspond to the standard deviation of the value added for the sector. Business cycle indicators are published quarterly, and the data for 2014 is based on the average of the first three quarters.

When analysing the data used by the Commission, a point worth noting is that the capacity utilisation rate series only extends to 1996. The worst crisis years of the 1990s recession, for

[11] The data from spring 2014 also provides a number of other explanatory variables, which I have use as auxiliary variables when assessing the Phillips curve. These consist of the change in terms of trade, which is estimated on the basis of the change in consumer prices and the GDP price ratio; the lagged change in terms of trade; the rate of change in labour productivity (GDP per number of workers); the acceleration of change in labour productivity; the lagged rate of change in labour productivity and the share of wages and salaries of GDP and its two lags. Their use, however, does not significantly affect the main results, and thus they are abstracted from the current paper (see, Kuusi, 2015, for further details.)

example, are therefore missing from the data. Therefore I also make use of another indicator series: estimates by industrial enterprises regarding their order books in relation to the norm, which I compiled by chaining indicator series BTEOLRSL and BTEOLL:B8S of the Confederation of Finnish Industries (EK). The data has been available since 1976; it thus includes data on the 1990s crisis. Further analysis provided by Kuusi (2015) suggests that the use of the alternative indicators yields very similar results.

Finally, government budget balance series is taken from the AMECO database in the spring of 2015.

3.2. Data used in the evaluations of the alternative measures

For a historical assessment of alternative indicators, I need information on revenue-related policy changes implemented in public finances (including central government, municipalities and social funds). With respect to central government finances, the data I have collected for this paper contains information on the estimated effects of changes in tax policy as provided by the Financial Status Reports 1977–2002. After the year 2002, the reports are no longer available in the same form. Therefore, I have evaluated the changes in the tax policy against the government's budget proposals for 2003–2008. With respect to the period 2009–2014, I received the necessary information from the Ministry of Finance. The Ministry's data also includes information on various types of deductions concerning the whole public sector. In addition to state taxation, I will examine the effects of policy changes made in general government finances. With respect to the period 2009–2014, I will use the evaluations of the Ministry of Finance. As for the preceding years, 1977–2008, I could not find direct estimates of the effects of changes made to the criteria for charges on revenues, so I used the observed changes in charge percentages as the basis for the effect estimates of the decisions.

I will evaluate local government finances' revenue estimates on the basis of changes in the weighted average local income tax rate and the real estate tax rate. I will calculate the euro-denominated effect of the change by multiplying the change in the tax base with the tax basis of the previous year, which in the case of local income tax means private income and in the case of real estate tax the taxable value of real estate. As for social insurance funds, I will evaluate the changes on the basis of the average social insurance contributions (employer's child benefit, accident, health, national pension, unemployment and TEL contributions and employee's unemployment and TEL contributions), expressed as percentages of the payroll. I will multiply the change in these with the previous year's total payroll.

The number of discretionary measures on the revenue side in my calculations corresponds fairly well to previous assessments (for more details, see Kuusi 2015). Perotti (2011) assessed discretionary total changes on the revenue side with regard to Finland during the crisis of the 1990s. The calculations that have now been completed reinforce the impression presented in the article that the revenue basis had a major impact on the overall balance of public finances during

the crisis. However, the results differ from the earlier evaluations by the IMF (see Perotti 2011), according to which public finances were not adjusted by increasing revenues but by cutting expenditure. In addition, the Commission's figures for 2010–2014 from the AMECO database (UDMGCR variable) are also parallel with the estimates used in this work.[12]

In addition to the evaluation of changes in the revenue basis, I have collected other variables needed for the calculation of alternative discretionary measures. Potential output growth estimates for 2011–2014 are based on reference values provided by the Commission to the individual member states. Potential output growth estimates for 2002–2010 are based on the estimates made by the Commission in the autumn of the same year, by applying the production function method. Potential output growth estimates for 1989–2001 are based on the estimates made by the OECD at the end of the same year on average growth for the following two years and the preceding five years. With respect to the 1980s, I have estimated potential output growth on the basis of the average five-year growth forecast made by ETLA (the Research Institute of the Finnish Economy) in the same year.

With respect to the expenditure benchmark, I will use the GDP inflation projections as inflation series. For the years 2001–2014, these are the European Commission's forecast averages from the previous year's spring and autumn. For the years prior to that, I will use the previous year's average inflation forecasts made by the Ministry of Finance. With respect to bottom-up evaluation, I will use the actual change in the GDP price.

For the other variables, I have followed the principle of trying to find the longest time series possible in order to enable a historical assessment. As expenditure series, G, I have selected a time series, published by the IMF, for general government total expenditure because this covers the longest period from the early 1980s onwards. In addition, I have used the Ministry of Social Affairs and Health's information on unemployment expenditure, which I will eliminate from the expenditure aggregate related to the bottom-up evaluation and, with respect to the expenditure benchmark, from the expenditure aggregate related to cyclical unemployment expenditure. As interest expenditure, I will use the time series given for property expenditure. The amount of public investment is based on the figures obtained from the National Accounts.[13]

The data on Finland's shares of EU structural funds is based on the data for 2010–2014 obtained from the audit memorandum prepared by the National Audit Office of Finland regarding compliance with the Stability and Growth Pact. Due to lack of preceding observations, I will set

[13] In order to enable comparability between the results, I will also use the alternative variables which the Commission applies in its assessments. From the AMECO database, I have collected series for general government expenditure (UUTGE), interest expenditure (UYIGE) and investments (UIGGO). However, expenditure aggregates cannot be calculated on the basis of these for the years before 1999.

these to zero prior to the year 2010. Likewise, I will not assess the amount of non-recurring items since the related evaluations are not available for the entire period in question. In any case, since they have also been eliminated from the output gap-based structural balance indicator presented by the Commission, they are not essential for comparison purposes.

4. Results

4.1. Evaluations of structural unemployment and total factor productivity

In the following, I will first examine the method for calculating structural unemployment. The short-term state of the labour market can be assessed with the help of the New-Keynesian Phillips curve, which describes the inverse relationship between (wage) inflation and unemployment. In principle, the connection to structural unemployment is clear. If inflation reacts to an increase in cyclical unemployment, the detected connection can be reversed, and the increase in cyclical unemployment can be specified efficiently with the help of inflation. Thereafter, structural unemployment can be achieved by removing the cyclical part from detected unemployment. In practice, However, price stickiness in major economic crises due to anchored inflation expectations or pressures not to lower wages have turned out to be problematic with regard to the assessment of cyclical unemployment (IFAC 2013; Wren-Lewis 2013; Krugman 2013). If they are not sufficiently taken into account in the models – or if the models do not identify them correctly – the result may be oversized assessments regarding the development of structural unemployment. Based on changes in inflation, an increase in unemployment can be considered structural, although it would in fact be cyclical. The output gap will be underestimated, as the increase in structural unemployment does not increase the output gap.

Explaining unemployment with the inflation indicator[14] used by the Commission is also problematic in Finland's case. There was no clear unambiguous connection between the variables, in particular during the crisis of the early 1990s (see the left panel in Figure 1). During the years of highest unemployment, strong inflation would have been required in order for such a connection to have been observed. This could not, however, be discerned on the basis of the data. The highest unemployment estimates were specifically for these years, based on the Commission's method (see the right panel in Figure 1).

In addition, special attention should be paid to the fact that the commission imposes inequality constraints on the parameter estimates when applying the estimation. In the Finnish case, the constraint restricts the maximum size of the cyclical change in unemployment forecasted by the model, insofar as the New-Keynesian Phillips curve does not directly explain it. Using the

[14] The inflation variable is a change in unit labour cost that is equal to wage inflation less the labour productivity growth rate and the change in consumer prices.

restrictions could lead us to underestimate the amount of cyclical unemployment (See, Appendix for further details). I recommend that the parametrisation of the method used for calculating structural unemployment be changed to better correspond to a plausible model based on the literature and observations outside the model. When the constraint used in the parameter estimation is removed, structural unemployment increases more moderately during the crisis of the early 1990s (see the right panel in Figure 1).[15]

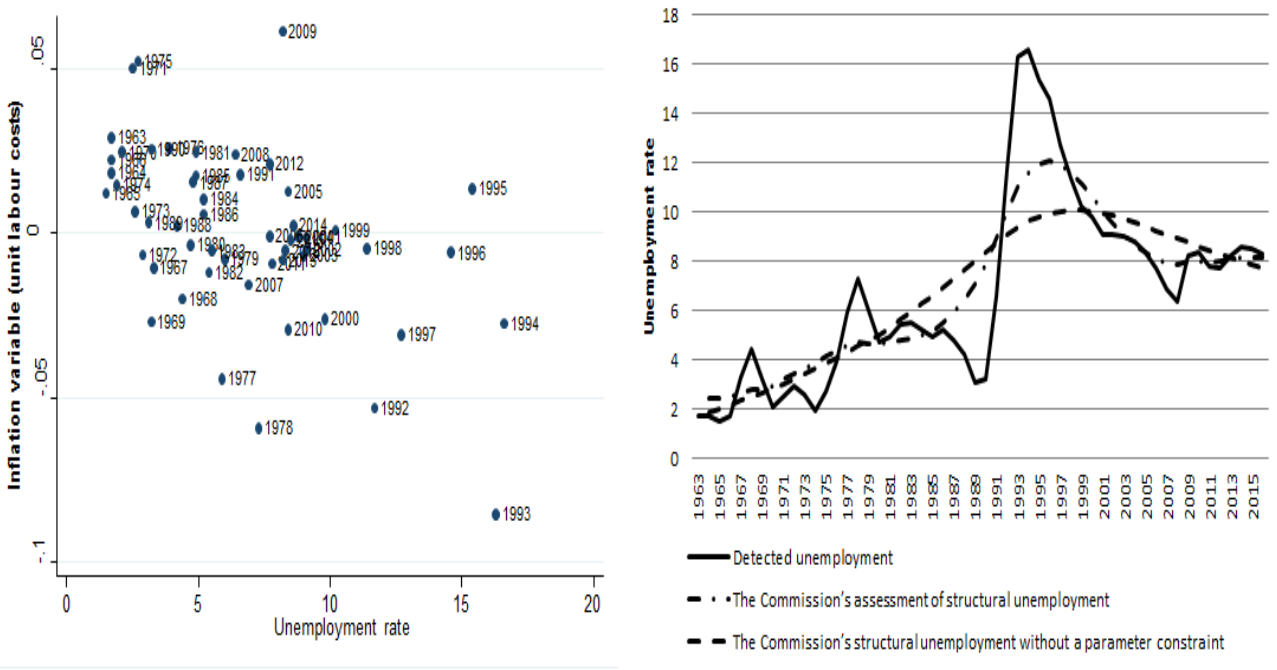

Figure 1 Source: Data and algorithms from the European Commission´s autumn 2014 forecast, and author's own calculations.

I have also examined the European Commission's assessments of structural total factor productivity. Figure 2 shows the natural algorithm of structural total productivity and an assessment of potential total factor productivity with the Commission's calculation method for 1980–2016. The dominant feature in the figure is the strong slowing of the total factor productivity growth rate after 2007. During the present crisis, the development of total factor productivity has been the main factor affecting potential output. For example, compared to the recession in the 1990s, the halt in total factor productivity growth has lasted significantly longer. Total productivity reached 1989 levels only a couple of years after the start of the crisis, whereas during the current crisis, total factor productivity was still far from the 2007 level in 2014.

While similar patterns are also seen elsewhere, reasons for the weak development of Finland's total factor productivity during the economic crisis have been searched for, particularly in the

[15] The development of Finland´s structural unemployment during the crisis of the 1990s has been assessed by Fregert and Pehkonen (2009), who also summarise the results of the previous literature. Their conclusion is consistent with the presented unconstrained model: the increase in structural unemployment would have been approximately 4 to 6 per cent during the crisis, and would have begun to decrease very slowly during the recovery phase.

industry-level shocks that have hit the economy. It has been argued that the fall in total factor productivity is due to problems in the Nokia-driven ICT cluster and in the paper and mechanical engineering industries. On the basis of productivity growth from sector to sector, and when Finland is compared to Sweden and the United States, it appears that the rather gloomy assessments about the development of total factor productivity that have been calculated using the Commission's method are fair. When the development of the total factor productivity is examined in various periods of time in Finland, Sweden and the United States, for example, it becomes evident that the growth rate of total factor productivity has been on average very similar in the said countries in 1995–2014. Following the crisis, the strong growth effect of ICT prior to the economic crisis is stabilising in the Nordic countries to the same level with the United States.

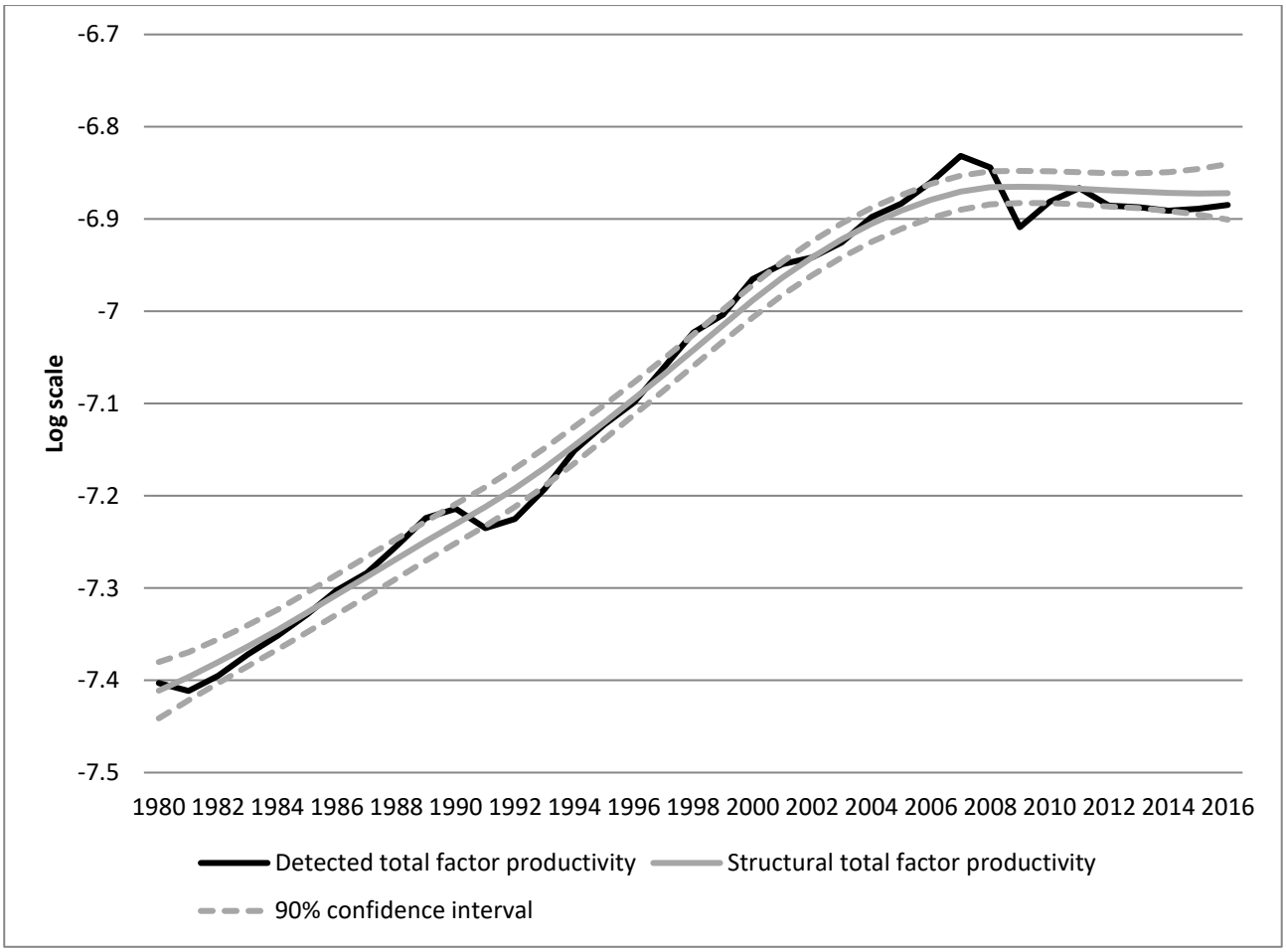

Figure 2 Data and algorithms from the European Commission´s autumn 2014 forecast, and author's own calculations.

4.2. Evaluations of the structural balance

When the gap estimates for different components have been calculated, they can be aggregated as an output gap in the economy. Measuring the structural budget balance used by the Commission is fairly straightforward. The estimated output gap is multiplied by cyclical elasticity (ϵ) and income is subtracted from the headline balance. I use the estimate of 0.57 provided by the Ministry of Finance in the spring of 2015 as the cyclical elasticity.

Figure 3 shows alternative structural balance estimates as well as a series of non-adjusted balance retrieved from the AMECO database. I have first calculated an ex post evaluation of a cyclical correction (an ex post evaluation of cyclically-adjusted structural balance) using the method recommended in the report Kuusi (2015), that is, I based structural unemployment on an assessment in which the above-mentioned parameter constraint has not been used.[16] In addition, I evaluated the operation of the indicator in (quasi) real time, without information on future development of the variables that are used in the estimation. I calculate the estimate after constraining the data being used at different points in time.[17] I adjust the output gap estimate that I recommended, which does not contain parameter constraints, for key turning years in the economic cycle (1989, 1993, 1997, 2001, 2003, 2007 and 2009) by changing the ex post estimate of total productivity and structural unemployment to real-time estimates (real-time cyclically-adjusted balance). I will make the adjustment by removing the difference between the ex post estimate and real-time estimate of both components from the output gap. At the same time, I do not comment on the real-time cyclical adjustment of other output gap components, such as the participation rate. The GDP and nominal deficit estimates are also ex post.

Real time has a considerable effect on the indicator's functioning. When estimates of total factor productivity and structural unemployment are based on data which takes no account of the trend for future years, the structural balance proves to be considerably more procyclical.[18] In real time, the structural balance has deviated materially from the ex post estimate in two of the three expansions in recent decades (1989, 2000, 2007). The structural balance is overestimated by approximately 1.3 percentage points with respect to three business cycle peaks, on average. In addition, the real-time structural balance underestimated the deficit component due to the economic crisis when the downturn of the early 1990s had already begun. For example, the 1993 ex post estimate of the structural contribution to the total deficit is approximately 35 per cent, while the real-time estimate would have been approximately 60 per cent.

The results suggest that the output gap method has limited capacity to predict cyclical changes in real time, and therefore, its use for steering fiscal policy could lead to a procyclical fiscal policy. On the basis of the figure, it seems that fiscal policy guided by an output gap-based SB does not react

[16] The suggested change in the calculation method of structural unemployment would have a positive effect of about 1 percentage point on the structural balance during the crisis of the 1990s. During the present crisis, the effect is not quite as great. For example, a change in the calculation method of structural unemployment in 2016 would have a positive effect of approximately 0.02 percentage points on the structural balance.

[17] To be precise, a genuine real-time analysis would require the selection – as data – of the time series actually in use during the year under scrutiny. As regards the unemployment series, the data is not revised ex post. However, later data or methodological changes may have influenced the inflation series. In addition, the Commission uses estimates for the next two years when measuring the structural deficit. More details on the real-time behaviour of the individual components can be found in Kuusi (2015).

[18] However, when examining the recession at the beginning of the 1990s, it can be observed that ex post estimates are rather procyclical, particularly when a crisis has emerged. The budget balance weakened by nearly 6 percentage points within a few years when the crisis broke out at the beginning of the 1990s.

in a contractionary manner during the economic upswing in the 1980s and early 2000s. On the contrary, the indicator permits a fiscal policy that is more expansive than the actual fiscal policy, if it is calculated without the future development of the economy that would be available later. Besides, an output gap-based SB indicator ignores the fairly strong contractionary measures in fiscal policy implemented in the crisis in the early 1990s.

It should be note that the real-time results presented are not without problems. Firstly, the real-time estimate of the present output gap may underestimate the accuracy of the Commission´s estimate, as the Commission uses forecasts of the trend for future years to support the estimate. If the forecasts are informative regarding cyclical change, they can improve the model´s accuracy. On the other hand, revisions may have taken place in the data used which are not taken into account by ex post cutting of the data. It should also be noted that the assessment of the real-time gap does not take into account the effect of changes in other output gap components (such as participation). Likewise, I will not assess the amount of one-off items since the related evaluations are not available for the entire period in question.

However, earlier literature would seem to indicate that there are no major differences between realised forecasts and quasi real-time assessments such as the one presented here. Kuusi (2014) compared quasi real-time output gaps with the Commission's genuine real-time estimates, and the results achieved with the method did not significantly deviate from each other. The average difference in the output gap estimates was about 1/2 a percentage point in 2006–2012, which corresponds to about a 1/4 percentage point effect on the structural deficit. Virkola (2013) also examined the Commission´s revisions in respect of 2007 and observed that real ex post revisions to the output gap in Finland were on the same scale as the estimates currently shown, i.e. approximately 5 percentage points. Since the one-off measures are eliminated from both the output-gap based structural balance and the discretionary fiscal policy measures, they are not essential for comparison purposes.[19]

[19] Koen and van den Noord (2005) show that the amount of one-off measures are likely to have been small during 1994-2000, with the exception of the postponement of tax refunds in 1994 (1.3 pps of GDP). Financial restructuring in the banking sector can explain only a minor fraction of the abysmal deteriation of public finances (-1.5 pps of GDP per annum in 1991-1992). In the 2000s, the one-off measures have been small according to the Ameco data.

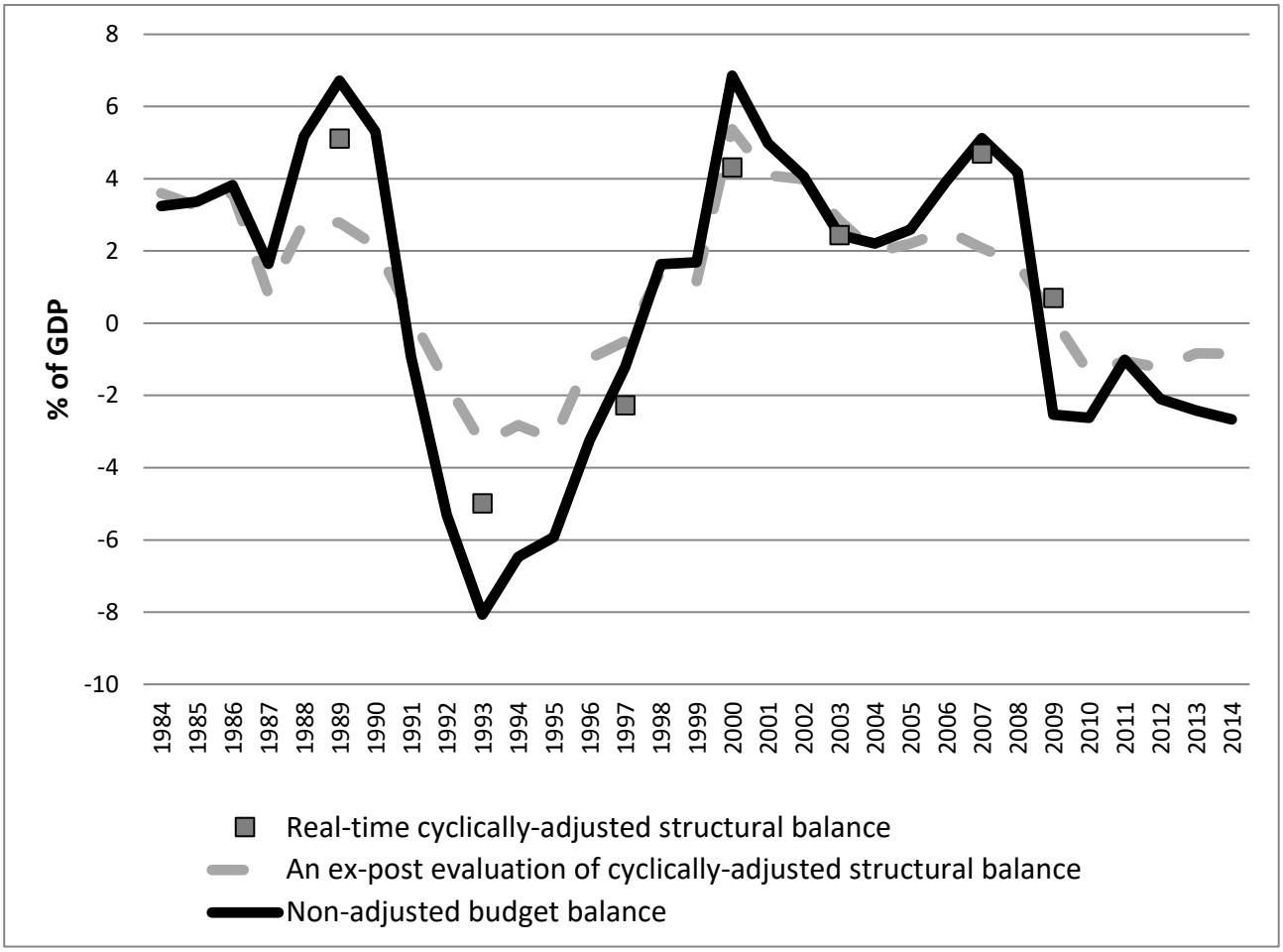

Figure 3 Data and algorithms from the European Commission´s autumn 2014 forecast, and author's own calculations.

4.3.The bottom-up alternatives

In Figure 4, I present assessments on the amount of discretionary fiscal efforts based on the bottom up method and the DFE indicator specified above. By applying the said method, an increase of one percentage point in the DFE indicator improves the structural balance by one percentage point. The cumulative change, on the other hand, indicates the total change in the budgetary position within a certain time period.

I will focus here on assessments according to the bottom up method, as they do not use ex post data on the development of the economy. In that way, the presented method offers a real-time baseline for the SB. For comparison, the figure contains the real-time cyclically-adjusted balance and the non-adjusted balance presented in the previous subsection. As discussed before, both indicators are estimated without the information on the one-off measures. Figure 4 also shows the DFE indicator based on the expenditure aggregate, which is calculated based on the expenditure rule. The adjustment items for different types of expenditure have a relatively minor effect on the

resulting interpretation of fiscal policy developments. On the other hand, the differences between the assessments are almost fully attributable to the used inflation variables.[20]

On the basis of Figure 4, a fiscal policy steered by discretionary measures could have become more countercyclical than a policy steered by SB. During the economic upswings in the 1980s and the 2000s, the fiscal policy is stimulative when measured using a discretionary assessment.[21] This observation enables (1) the fiscal policy guided by the method to be tighter during the upswings (2) the control of the overheating of the economy and (3) creation of a margin for recovery measures during the crisis. In comparison, before the outbreak of each of the two major crises, the real-time structural balance based on the output gap method was exceptionally strong, which could have enabled the continuation of a stimulative fiscal policy during the upturns.

On the other hand, the tightening of the fiscal policy after the outbreak of the crisis in the 1990s is clearly visible on the basis of the discretionary indicator. In particular, the significant tightening on the revenue side of social insurance contributions explains the strong increase in the discretionary indicator. After the outbreak of the crisis, the measured fiscal policy was tightened rapidly from 1992 onwards and continued throughout the 1990s. The observation could have enabled the use of a more stimulative fiscal policy. When comparing the results to the development of the output gap-based structural balance (see Figure 3), we can see that, on the basis of the latter, the tightening of the fiscal policy did not begin until after the mid-1990s even as an ex post evaluation.

[20] The inflation variables applied make the assessments somewhat cyclical, although the inflation projections used for the expenditure benchmark reduce the effect of inflation somewhat, particularly with regard to the end of the 1980s. Both indicators allow for strong growth in expenditure during periods of high inflation, while during crises and periods of low inflation the need may arise to make additional cuts in public expenditure. On the other hand, taking account of inflation adjustments in income tax rates as a change in the revenue basis has, to some extent, the opposite effect.

[21] Further decomposition of the indicator into revenue and expenditure components (DFE_t^R, DFE_t^E) suggests that in the 1980s the weakening of the fiscal position was due to increases in public expenditures, whereas in the 2000s it was due to both higher expenditures and lower revenues.

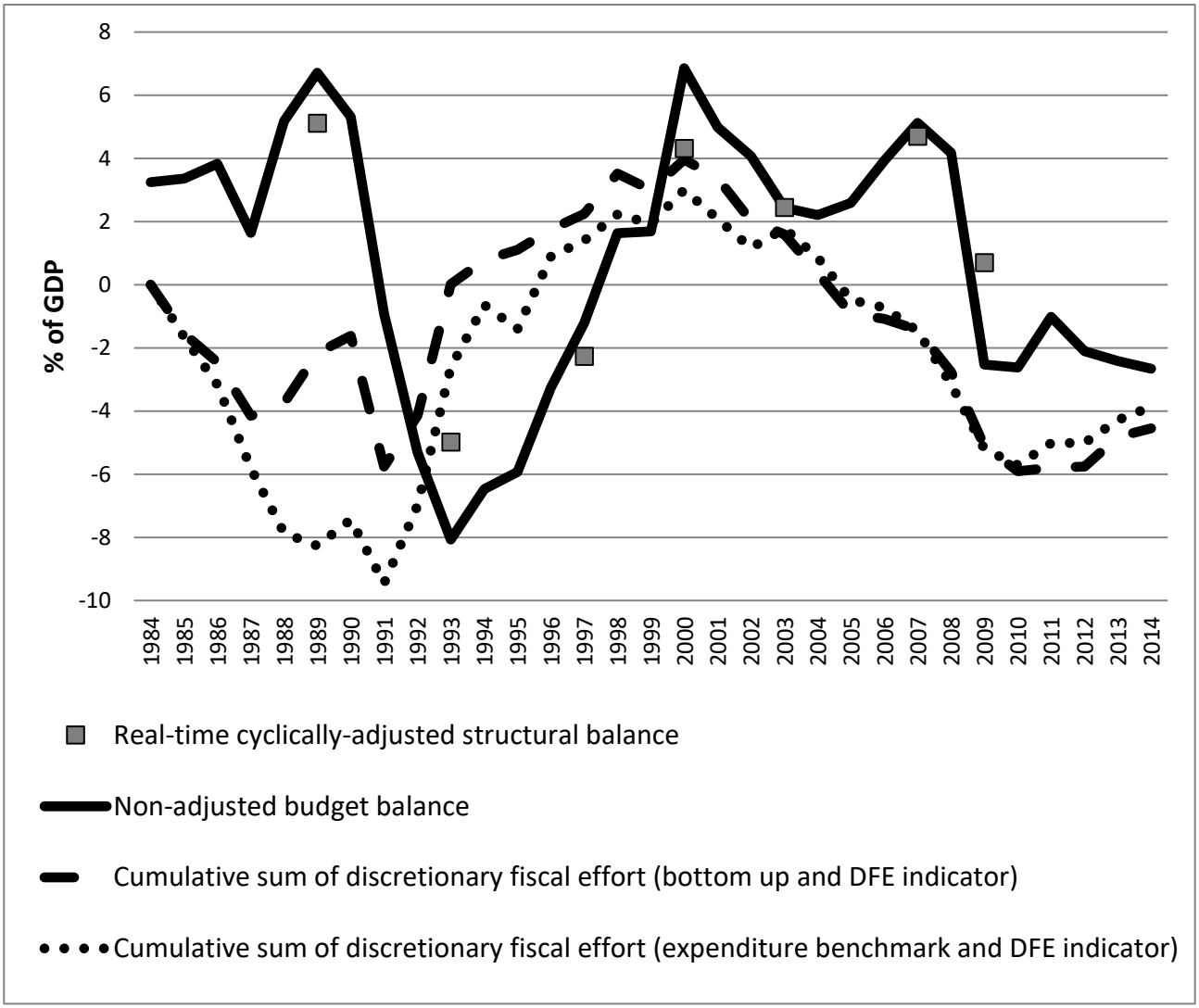

Figure 4

Measured by both indicators, the fiscal policy was stimulative at the initial stage of the present crisis, but from 2011 onwards, the indicators diverge again as the discretionary fiscal effort indicator suggests a 2–3 percentage point tightening of the fiscal policy in 2010–2014, whereas the ex post structural balance indicator (in Figure 3) shows hardly any signs of improved public finances.

In order to further illustrate the difference between the indicators in terms of cyclical behaviour, I have compared the ex-post changes in the output gap to the change in various fiscal policy indicators in 84–89, 89–93, 93–97, 97–00, 00–03, 03–07 and 07–09.[22] Based on a cross-tabulation, a one percentage-point growth in the (ex post) output gap would have weakened the real-time structural balance by approximately 0.65 percentage points. Meanwhile, the bottom-up assessment method does not indicate a clear connection between the cyclical change and indicator developments.

[22] The output gap is measured using the Commission's method.

What explains the differences between the methods? First, similarly to the European Commission (2013B) and Carnot and de Castro (2015), I compare the developments in the structural balance net of interest expenses with the bottom-up method. After netting interest expenses, the differences between the indicators are more clearly attributable to methodological factors, such as different cyclical adjustments of revenue and expenditure items and a different method of calculating the potential output growth rate. The comparison between the real-time cyclically-adjusted structural balance and the structural primary balance suggest that the interest expenses explain around 15 per cent of the cyclical changes in the real-time structural balance, while the rest is explained by methodological differences.

Second, by means of changes in the revenue basis, the cyclical development of revenue items can be directly examined in different years, instead of using the output gap and fixed cyclical elasticity. I will do this by first eliminating the discretionary changes related to the various income types presented above (for methodological details, see, Kuusi 2015, appendix 4; Carnot and De Castro 2015). In principle, the remaining element of income development can be evaluated as a (cyclical) change independent of fiscal policy, while naturally taking account of any errors in the effects of the related decisions. The results of the analysis suggest that changes in cyclical income in relation to cyclical income evaluated using the fixed elasticities may explain various percentage points of the differences in the changes in the structural balance according to the different indicators. In particular, the analysis of the late 1980s and early 2000s suggests that the income growth experienced during the economic upturn exceeded the estimates produced using the output gap method. In downturns, on the other hand, it appears that the differences are attributable to different cyclical adjustments of the unemployment expenditure and the different method of calculating the potential output growth rate.

4.4.Indicator differences from the viewpoint of EU regulations

I have also evaluated how the indicators would have steered the financial policy in practice if EU's fiscal policy rules would have been followed during the period of time. I collect the observations in the previous analysis, and examine the regulatory requirements that the country would have faced in different time periods. The rules are based on a summary of the criteria applied regarding deviations from the rules of the preventive and corrective arms of the SGP (European Commission 2013A, 2016):

- Deviation in the preventive arm:
    - The deviation from the medium-term objective (MTO)[23], was more than 0.25 per cent of the GDP in the previous year (ex post assessment) **and**

[23] The preventive arm of the SGP (applied when the 3 and 60 per cent deficit and debt criteria of the excessive deficit procedure criteria are not breached) uses the structural budgetary position and the increase in expenditure to assess deviations from the MTO or from the path towards it. The Fiscal Compact obliges the member states to set an MTO; as a result, this obligation is included in their national legislation. In the Fiscal Compact, the lower limit of the MTO for countries in the Eurozone was set to a budgetary position of -0.5 percent, except in the case of countries whose debt is

    - the nominal deficit does not exceed 3 per cent, i.e. the country is not subject to the corrective arm of the SGP[24] **and**
    - on the path towards the MTO, the structural balance improves by less than 0.5 percentage points **and**
    - the structural deviation from the path (ex post) is significant, i.e. at least 0.5 per cent of GDP **and**
    - the deviation is significant from the viewpoint of both the structural balance and the expenditure benchmark while taking account of the cyclical state in accordance with the guidelines of the European Commission (2015, appendix 2).
- Deviation in the corrective arm:
    - The nominal deficit exceeds 3 per cent **and**
    - the measures are not effective, i.e. the country is unable to adjust its structural budgetary position by at least 0.5 percentage points (structural balance adjustment path) **and**
    - the structural deviation from the path (ex post) is significant, i.e. at least 0.5 percentage points per year **and**
    - the deviation is significant in terms of both the structural balance and the bottom-up assessment.

The analysis shows that in many years during economic upturns (1985–1988, 2004–2005, and 2007–2009) Finland would have achieved the MTO based on the real-time structural balance, but at the same time would not have achieved the expenditure growth rate required by the expenditure benchmark, or would have been close to exceeding it. The major strengthening in the structural balance that preceded the crisis of the early 1990s and the current crisis could have allowed an expansion in the public finances. During both periods, the structural balance was rather strong as measured on the basis of both ex post and real time output gap estimates. At the same time, the expenditure benchmark might have imposed stricter limits on fiscal policy during the said years.

Between the years 1993–1996, during which the legislation related to the corrective arm of the SGP could have been applied on the basis of the deficit criterion, Finland would not have reached the 0.5 percentage point adjustment requirement in the crisis years 1993 and 1995. Thus, Finland would have been unable to sufficiently adjust its structural deficit, and further measures might have been required. However, following a careful consideration based on the bottom-up indicator, it can be seen that a sufficient adjustment of the general government balance was implemented in those years. [25]

4.5.Discussion of the results

---

less than 60% and which do not have long-term sustainability risks (in which case the lower limit is -1 per cent). In the following analysis, the MTO is set at -0.5 pps of GDP.

[24] Here, I abstract from the debt rule, as the debt level of the Finnish economy has remained below the 60% per GDP benchmark during the whole time period.

[25] In the present situation the differences between the indicators are smaller. An output gap-based structural balance would threaten to breach the limits of the EU's fiscal policy rules, whereas according to of the expenditure benchmark, tightening in fiscal policy has been sufficient to compensate for the negative effect of downturn on the government finances.

All in all, the empirical analysis of the rules reinforces the impression of the limited capacity of the output gap method to predict cyclical changes in real time. On the other hand, the discretionary indicators can provide better guidance and steer fiscal policy more counter cyclically. In particular, it seems that the fiscal policy based on the expenditure rule would have been contractionary during the lead up to the 1990s crisis, which could have helped to alleviate the crisis and increase the margin for recovery measures after it started. On the other hand, based on a discretionary bottom up assessment, the contractionary fiscal policy practised from 1992 onwards would have been sufficient, and unlike the output gap-based SB, the method would not have generated additional contractionary pressures. [26]

The historical analysis raises several concerns regarding the use of the current output-gap methodology. It seems fair to say that explaining unemployment on the basis of inflation may be problematic, especially in the case of major economic crises, like the Finnish Great Depression of the 1990s. In particular, the findings are consistent with the recent literature suggesting that the behavior of inflation does not necessarily correspond to the New-Keynesian Phillips curve during major crises, even though it includes a delayed inflation term. Moreover, the output-gap revisions are not just due to methodological issues, but they reflect genuine real-time uncertainty regarding the measurement of potential output. The measured revisions during the Finnish Great Depression are large just as they have been large during the current crisis.

While they seem to work better, the expenditure benchmark and bottom-up assessment are not fully immune to measurement problems either. They should also be buttressed by an understanding of the medium-term output potential of the economy. Although the moving average for past trends and forecasts over the business cycle is less sensitive to cyclical changes, short-term positions may also be reflected in longer-term assessments. Currently, this is a particularly important problem, as many countries are facing uncertainty regarding their long-term productivity prospects. Furthermore, independent economic analysis of the effects of various changes in policy is needed to back up the expenditure benchmark and the bottom-up approach. For example, any appraisals of the magnitude of the multiplier effect of fiscal policy – both during and outside crises – remain fairly contradictory. In terms of finding the neutral policy stance, the inflation variable of the expenditure benchmark and the bottom-up approach should be replaced with longer-term equilibrium inflation, in order to avoid changes in inflation or its forecasts having the effect of enhancing cyclicality. Both indicators allow for strong growth in expenditure during periods of high inflation, while during crises and periods of low inflation the need may arise to make additional cuts in public expenditure.

[26] Based on the earlier literature it seems that the countercyclical steering of the fiscal policy may be beneficial also in terms of maximizing the magnitude of the positive multiplier effect of fiscal policy. Keränen and Kuusi (2016) estimates a regime-switching STVAR model by using the Finnish data of the discussed time period. They show that the consolidation measures may slow the economic recovery especially at the turning point of the business cycle in the early 1990s, while during the economic upturns the multiplier effects may be small, or even negative.

Finally, based on the functioning of the historical behavior of the SB, it is problematic that the expenditure benchmark plays no clear role – independent of the structural balance and its calculation methods – in the EU´s fiscal rules. In determining the medium-term growth reference rate of potential output in accordance with the expenditure benchmark, the preventive arm of the SGP still relies on the fulfilment of the MTO. In fiscal policy legislation, if the MTO has been achieved in a certain year, the reference growth rate of expenditure is the long-term GDP growth. On the other hand, if the MTO has not been achieved, expenditure growth measured using indicators must be slower so that the deficit decreases by at least 0.5 percentage points per year.[27] This link is necessary, because the discretionary measures as such do not involve monitoring the objective level of fiscal policy, but only changes in fiscal policy. An alternative solution could involve tying the expenditure benchmark more closely to the debt level and to forecasts of its future trends based on sustainability calculations. Hughes Hallett and Jensen (2012), for example, propose a given limit for the indebtedness level below a GDP ratio of 60%, where exceeding such a limit would trigger preventive measures. Although the debt ratio is also sensitive to cyclical changes, it is not as prone to fluctuation as the (structural) deficit. On the other hand, we already have experience of cyclical adjustments of debt under the present rules.

5. Conclusions

The structural balance has played a central role in the EU's new fiscal policy legislation. In this study, I examine estimates of the structural balance from a historical perspective, using the European Commission's analysis method and comparing it to alternative fiscal policy indicators. The perspective adopted is that of recent Finnish economic developments between the years 1984-2014; a time period that includes the Finnish Great Depression of the 1990s.

The results of the study corroborate the view presented in earlier literature, according to which a structural balance is difficult to estimate using the output gap approach. Although the European Commission uses the latest statistical methods to assess the cyclical state of the economy, measuring the output gap in real time proves to be a difficult task in practice. The capacity of the output gap method to filter out cyclical fluctuations and measure cyclical phase effects on the budgetary position is limited, which may result in an under- or overestimate of the structural budgetary position. Particularly in the overheating phase of the Finnish economy in the 1980s and also in the deep economic crisis of the early 1990s, using the structural balance to steer fiscal policy results in a very procyclical policy.

The results also indicate that discretionary indicators (expenditure benchmark and bottom up assessment) form a useful part of the EU's control of fiscal policy. On the basis of the analysis, fiscal policy based on the discretionary measures is more countercyclical compared to the policy

[27] Furthermore, the measurement of the expenditure rule builds on the output gap-based assessments of cyclical unemployment.

that is guided by the output gap-based SB. Having said that, established ways of analysing the slackness in the economy with the help of inflation in particular should not be ignored, although methodological challenges have been encountered in applying them in practice. Long-term growth estimates calculated based on discretionary methods do not offer a clear reference point for assessing the cyclical position of the national economy. However, it is necessary to develop the output gap method in this respect in order to improve its reliability.

Appendix: Parameter restrictions used when estimating the structural unemployment

With respect to the application of the model (that is specified more closely below), I paid particular attention to the parameter constraints used when determining the model. When considering the details of the model´s solution, it is observed that the parameter constraints concern the variance of cyclical variation (V_c), which is restricted, bound by an assumption, to a value of 0.5. Likewise, the shock variance (V_p) directly affecting the trend is restricted, in accordance with an assumption, to zero. Although the latter assumption is a natural variance positivity requirement, there are no clear statistical or theoretical grounds for the first assumption.

The constrained parameter concerns the size of the variance of the cyclical component of unemployment in so far as the New Keynesian Phillips curve does not directly explain it. The smaller the parameter, the more the estimate on the size of cyclical unemployment is based on inflation changes. Hence, the effects of parameter constraint depend on the function of the Phillips curve. If wage inflation clearly reacts to growth in unemployment, the observed connection can be turned around, and unemployment growth can, for its part, be effectively determined based on inflation. The analysis suggests that the impact of the constraint is large. Furthermore, Kuusi (2015) shows that there are differences between countries in terms of the constraint's size. [28]

Specification of the model

[28] It is worth noticing that in assessing which model is superior, the European Commission also draws particular attention to a few technical details with respect to which there is no clear difference between the models. Firstly, the dependence between the unemployment gap and the cost variable should be significant. This condition is fulfilled in both the unconstrained (the restriction is omitted) and the constrained model used by the Commission (let us denote them as models 1 and 2, respectively). In addition, the predicted cyclical component of unemployment should correlate negatively with the component explained by means of the inflation model in such a way that, within the model, the unemployment gap creates pressure to lower price levels. In both models, the correlation is negative (-0.58 in the model 2). Moreover, the model´s coefficient of determination for changes in inflation should be as high as possible. In neither model (Model 1 used by the Commission nor the unconstrained model, 2) does the coefficient of determination rise very high. When inflation is explained based on the component predicted by the model, the coefficient of determination is R2 = 0.2 in the constrained model, 1, used by the Commission, whereas the coefficient of determination used in the unconstrained model, 2, is R2 = 0.16. Finally, in terms of the selection of the trend model assumption (equations 11 and 12) in the case of the alternative model, 2, both autoregressive terms of trend pt are significant, confirming that the second order random walk is an appropriate assumption.

The functional forms and assumptions are the same as the assumptions concerning Finland in the forecast for autumn 2014. The first equation of the model is a regression model with structural time series error terms

$$u_t = \sum_{i=1}^{M_1} \alpha_{1i} z_{1it} + x_t, \quad (1)$$

where z_{1t} are exogenous variables. The error term x_t is a sum of the trend component p_t and cyclic component c_t so that

$$x_t = p_t + c_t. \quad (2)$$

The cyclical component is defined as the AR(1) model:

$$c_t = \phi_{c1} c_{t-1} + a_{ct}, \quad (3)$$

where a_{ct} is a cyclical shock term with a variance of V_c.

With regards to Finland, a trend shock is modelled in the reviewed Commission's method as a second order random walk defined by the following equations

$$p_t - p_{t-1} = \mu_{t-1} + a_{pt} \quad (4)$$

$$\mu_t - \mu_{t-1} = a_{\mu t} \quad (5)$$

The equations a_{pt} contain a shock that affects trends directly and has a white noise distribution. Its variance is V_p. The second shock $a_{\mu t}$ affects the slope of the trend and is also white noise. Its variance is marked as V_μ.

Another equation used in the method is the New-Keynesian Phillips curve that can be expressed concisely as:

$$RULC_t^w = \alpha + \gamma RULC_{t-1}^w + \psi_0 c_t + \psi_1 c_{t-1} + e_t, \quad (6)$$

where $RULC_t$ is the rate of change or real unit labour costs, c_t is the cyclical component of unemployment and α is a term that includes various long-term relations (such as the average rate of the increase in productivity). e_t is a noise term.

The difference compared to the traditional method of estimating a Phillips curve lies in the assumption that the cyclical component c_t cannot be perceived. Instead, the problem with a

statistical method of calculation is the assessment of unperceivable variables p_{tf} ja c_t using the maximum likelihood method and the so-called Kalman filter. A more detailed description of the method is presented by Planas and Rossi (2004), and Planas and Rossi (2014), These papers contain descriptions of the closed form solution for an estimator and of deriving the confidence intervals for forecast errors.

Literature

Barrios, S. and Fargnoli, R. (2010). "Discretionary measures and revenues in the run-up to the financial crisis", European Commission Economic Papers 419.

Begg, I. (2014) "Fiscal policy transparency" in Forssbaeck, J. and Oxelheim, L. eds. The Oxford Handbook of Economic and Institutional Transparency, Oxford: OUP

Begg, I. (2016). Fiscal and other rules in EU economic governance: helpful, largely irrelevant or unenforceable?, The Horizon 2020 programme FIRSTRUN project deliverable 6.4

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

Bergman, U. M., Hutchison, M. M., and Jensen, S. E. H. (2016) "Promoting sustainable public finances in the European Union: The role of fiscal rules and government efficiency" European Journal of Political Economy Volume 44, September 2016, Pages 1–19

Besley, T. (2007) "Principled Agents? The Political Economy of Good Government." Oxford University Press

Bouis, R., Cournède, B. and Christensen, A. (2012) "Implications of Output Gap Uncertainty in Times of Crisis", OECD Economics Department Working Papers 977.

Bryson, A. and Forth, J. (2015) "The UK's Productivity Puzzle", NIESR Discussion Paper No. 448

Carnot, N. and de Castro, F. (2015) "The Discretionary Fiscal Effort: An Assessment of Fiscal Policy and its Output Effect", European Economy Economic Papers 543.

D'Auria, F., Denis, C., Havik, K., Mc Morrow, K., Planas, C., Raciborski, R., Röger, W., Rossi, A. (2010) The Production Function Methodology for Calculating Potential Growth Rates & Output Gaps. European Economy, Economic Papers, 420.

Daly, M. C. and Hobijn, B. (2013) "Downward Nominal Wage Rigidities Bend the Phillips Curve." Federal Reserve Bank of San Francisco Working Paper 8.

Eschenbach, F. and Schuknecht, L. (2002) “Fiscal balances and asset prices”, ECB Working Paper 141.

European Commission (2006) “Public finances in EMU”, European Economy 3|2006.

European Commission (2013A) “Vade mecum on the Stability and Growth Pact”, European Economy Occasional Papers, Issue 151.

European Commission (2013B) “Report on Public Finances in EMU 2013”, European Economy 4|2013.

European Commission (2015) “Making The Best Use Of The Flexibility Within The Existing Rules Of The Stability And Growth Pact”, Communication From the Commission to the EP, the Council, the ECB, the ESC, the CR And the EIB.

European Commission (2016) “Vade mecum on the Stability and Growth Pact. 2016 edition”, European Economy Institutional Papers, Issue 021.

Fernald, J. G. (2014) “Productivity and Potential Output Before, During, and After the Great Recession”. Federal Reserve Bank of San Francisco Working Papers 2014-15.

Fregert, K. and Pehkonen, J. (2009) “The crisis of the 1990s and unemployment in Finland and Sweden”, in Jonung, L., Kiander, J., Vartia, P. (ed.) *The Great Financial Crisis in Finland and Sweden*, Edward Elgar, pp. 158-183.

Golinelli, R. M. S. (2008) “The cyclical response of fiscal policies in the euro area. Why do results of empirical research differ so strongly?”, Banca d’Italia Working Papers 654.

Gorodnichenko, Y., Mendoza, E. G. and Tesar, L. L. (2012) “The Finnish Great Depression: From Russia with Love.” American Economic Review, vol. 102, issue 4, pp. 1619–44.

Hallerberg, M., Strauch, R., von Hagen, J. (2007) “The design of fiscal rules and forms of governance in European Union countries”, European Journal of Political Economy vol. 23, issue 2, pp. 338-50

Havik, K. et al. (2014) “The Production Function Methodology for Calculating Potential Growth Rates and Output Gaps”, European Economy Economic Papers 535.

Hetemäki, M. (2015) “Eurokriisin syyt ja euroalueen tulevaisuus [Reasons for the Euro crisis and the future of the Eurozone]” *Kansantaloudellinen aikakauskirja* 1.

Hughes Hallett, A. and Jensen, S. E. H. (2012) “Fiscal Governance in the Euro Area: Institutions vs. Rules.” Journal of European Public Policy, vol. 19, issue 5.

IFAC (2013) “Assessment of Compliance with Fiscal Rules”, Irish Fiscal Advisory Council Fiscal Assessment Report 1.

Jalava, J., Pohjola, M., Ripatti, A. and Vilmunen, J. (2006). ”Biased Technical Change and Capital-Labour Substitution in Finland, 1902–2003”. *B.E. Journal of Macroeconomics*, vol. 6, issue 1.

Keränen, H., Kuusi, T. (2016) “The EU’s Fiscal Targets and Their Economic Impact in Finland”. ETLA Working Papers No 33.

Koen, V. and van den Noord, P. (2005) “Fiscal Gimmickry in Europe: One-Off Measures and Creative Accounting”, OECD Economics Department Working Papers, No. 417.

Kremer, J. et al. (2006) “A Disaggregated Framework for the Analysis of Structural Developments in Public Finances”, ECB Working Paper 579.

Krugman, P. (2013) “The Pain in Spain is not hard to explain”, Blog post.

Kuusi, T. (2013) “Structural Change in the Information Era: Perspectives from a Multi-Sector model of the Finnish Economy.” *Dissertation for Aalto University School of Business.*

Kuusi, T. (2014) “Consolidation under the Europe’s New Fiscal Rules: Analyzing the Implied Minimum Fiscal Effort”. ETLA Working Papers No 23.

Kuusi, T. (2015) “Alternatives for Measuring Structural Budgetary Position”, *Publication series of the Government’s analysis, assessment and research activities* 5/2015.

Lane, P. (2012) “The European Sovereign Debt Crisis”, *Journal of Economic Perspectives* 26, pp. 49-68.

Luoma, A. and Luoto, J. (2010) “The Aggregate Production Function of the Finnish Economy in the Twentieth Century”, *Southern Economic Journal*, 76(3).

Marcellino, M. and Musso, A. (2010) “Real Time Estimates of the Euro Area Output Gap – Reliability and Forecasting Performance”, ECB Working Papers 1157.

Morris, R. S. L. (2007) “Structural Balances and Revenue Windfalls – The Role of Asset Prices Revisited”, ECB Working Paper Series 737.

Mourre, G., Isbasoiu, G.-M., Paternoster, D. and Salto, M. (2013) “The cyclically-adjusted budget balance used in the EU fiscal framework: an update.”, European Economy Economic Papers 478.

Murray, J. (2014) “Output gap measurement: judgement and uncertainty.” UK Office of Budget Responsibility Working paper 5.

Orphanides, A. and van Norden, S. (2002) “The Unreliability of Output gap Estimates in Real Time”, *The Review of Economics and Statistics* 84(4).

Perotti, R. (2011) “The ‘Austerity Myth’: Gain Without Pain?”, BIS working papers No 362.

Planas, C. and Rossi, A. (2004) “Can Inflation Data Improve Real-Time Reliability of the Output Gap?”, *Journal of Applied Econometrics*, 19(1).

Planas, C. and Rossi, A. (2014) “Program GAP: Technical Description and User-manual, version 4.3”, Joint Research Centre Technical Reports, European Commission.

Pohjola, M. (2011) “Kasvukauden tilinpäätös” [Account of a period of growth], *Kansantaloudellinen aikakauskirja* 3.

Portes, J., Wren-Lewis, S. (2015) "Issues in the Design of Fiscal Policy Rules," Manchester School, University of Manchester, vol. 83, pages 56-86, 09.

Price, R. and Dang, T.-T. (2011) “Adjusting Fiscal Balances for Asset Price Cycles”, OECD Economics Department Working Papers 868.

Rünstler, G. (2002) “The Information Content of Real-Time Output Gap Estimates: An Application to the Euro Area”, ECB Working Paper 182.

Sacchi, A. and Salotti, S. (2015) “The impact of national fiscal rules on the stabilisation function of fiscal policy”, European Journal of Political Economy, 2015, vol. 37, issue C, pages 1-20

Schackleton, R. (2013) “Total Factor Productivity Growth in Historical Perspective”, Congressional Budget office Working Paper 2013-01.

Stock, J. H. and Watson, M. W. (2010). “Modeling Inflation after the Crisis. Economic Policy Symposium”. Federal Reserve Bank of Kansas City.

Summers, L. H. 2014. “Reflections on the ‘New Secular Stagnation Hypothesis.’” In Teulings, C. and Baldwin, R. (2014) “Secular Stagnation: Facts, Causes, and Cures”, London: CEPR Press.

Virkola, T. (2013) “Rakenteellinen budjettitasapaino” [Structural budget balance], Reports of the Research Institute of the Finnish Economy ETLA, No. 11.

Virkola, T. (2014) “Real-Time Measures of Output Gap and Fiscal Policy Stance”, Reports of the Research Institute of the Finnish Economy ETLA, No. 37.

Wolswijk, G. (2007) “Short- and long-run elasticities. The case of the Netherlands”, ECB Working paper 763.

Wren-Lewis, S. (2013) “How not to run fiscal policy: more lessons from the Eurozone” Blog post (Mainly Macro, 17 October 2013).