Fiscal Policy and the Shifting Goalposts

Antonio Fatás
INSEAD, CEPR and ABFER

This version: February 28, 2018

Abstract: This paper studies the negative loop created by the interaction between pessimistic estimates of potential output and the effects of fiscal policy during the 2008-2014 period in Europe. The crisis of 2008 created an overly pessimistic view on potential output among policy makers that led to a large adjustment in fiscal policy during the years that followed. Contractionary fiscal policy, via hysteresis effects, caused a reduction in potential output that not only validated the original pessimistic forecasts, but also led to a second round of fiscal consolidation. This succession of contractionary fiscal policies was likely self-defeating for many European countries. The negative effects on GDP caused more damage to the sustainability of debt than the benefits of the budgetary adjustments. The paper concludes by discussing alternative frameworks for fiscal policy that could potentially avoid this negative loop in future crises.

This paper was prepared for the The Euro at 20 conference organized by the IMF and Central Bank of Ireland in Dublin, June 25-26, 2018.

1. Introduction.

We are approaching the 20-year anniversary of the Euro area and the debate is still open about whether such a diverse group of countries form an optimal currency area. In the absence of national monetary policies, how can countries manage asymmetric shocks when labor mobility is limited? The main policy lever is fiscal policy. What have the last 20 years taught us about the EU fiscal policy framework? Did it perform as expected supporting national economies to absorb asymmetric events?

We asked similar questions 10 years ago when celebrating the first decade of the Euro. In the Fall of 2008 I was asked to participate in a book reviewing the first 10 years of the Euro, and the overall conclusion on fiscal policy was positive (Fatas and Mihov (2009)). The Euro had been a success and defeated the pessimistic scenarios some had predicted. But, of course, the economic conditions of the first nine years of the Euro were benign, years of relative calm from a macroeconomic point of view except for a small and synchronized crisis in the years 2002-2003.

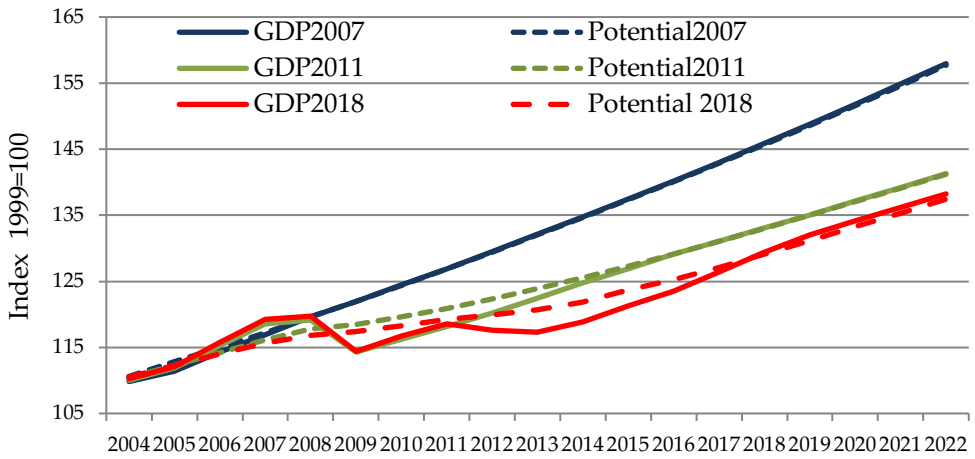
Despite the optimistic reading of the early years, there was no time to call victory and there were clear potential weaknesses and risks ahead. First, fiscal policy remained procyclical in the Euro, more so than in other countries (for example, the US). Second, the fiscal policy rules originally defined by the Maastricht Treaty had failed to provide a credible long-term framework for sustainability and continued to be ill-designed to handle a large asymmetric shock.

The conclusion was that: *"one cannot ignore that these results also portray a picture of a failing institutional framework for fiscal policy. One that has been focused on the concepts of sustainability and coordination without much success, at a time where it should have been focused on* ***strengthening automatic stabilizers and designing a proper framework for the conduct of fiscal policy over the business cycle****."*

Unfortunately, this warning and its associated risks became evident as the book went to press (January 2009). A large economic and financial crisis starting in the Fall of 2008 put the system to a test. A test that it failed through a combination of monetary policy hitting the zero-lower bound, the ECB not realizing early of its implications and fiscal policy, in particular after 2010, becoming highly procyclical because of the reaction to the increase in government debt levels.

The outcome of this failing policies can be summarized in the behavior of the Euro economy during these years. In Figure 1 we show the projections for both GDP and potential output for the Euro area in three different dates: April 2007 (before the crisis), April 2011 (after the first wave of the crisis) and April 2018, the latest available data.[1] GDP forecasts where downgraded several times.

Figure 1. Revisions to Euro actual and potential GDP



The revisions were large in the first years, but they continued even after 2012 as the Euro zone entered its second recession. And when forecasts were being revised downwards, they also did so for long horizons.

What is clear from this chart is that the effects of the crisis were very large and persisted over time. Relative to the trend that the Euro area was following prior to the Euro launch in 1999, GDP today is still far below that level. The IMF expects today that by 2022 the Euro area will be about 15% below the level implied by its pre-crisis trend.

What was the role of fiscal policy during these years? Contractionary fiscal policies were adopted two years after the crisis had started. The reason for this was that several Euro economies saw exploding deficits and debt in the early

[1] The April 2007 WEO does not contain forecast beyond 2008 for GDP or Potential. In that case, we are we are extrapolating both series using the average growth rate since 1999. The April 2011 WEO contains forecasts up to 2016. We are extrapolating both series for the next six years using the average growth rate since 1999. The April 2018 WEO contains forecasts up to 2022 for both variables. GDP data prior to 2007 is not identical in all three vintages because of data revisions. Potential was also revised backwards for several of these years.

years of the crisis. The conclusion was that a process of fiscal consolidation was necessary, partly to respond to new market conditions (which in some cases had stopped funding of government debt), partly as a way to restore confidence and growth. While the debate on what happened is (and will likely continue to be) open, there is strong evidence that fiscal policy played a big role in those years and GDP growth suffered as a result (Blanchard and Leigh (2013)).

But the effects of fiscal policy in output might have then fed into pessimistic views on the future and triggered additional fiscal consolidations and be partly responsible for the next recession. This negative loop happened through revisions to estimates of long-term GDP or potential output. When designing fiscal policy, in particular when sustainability is an issue, governments need to have an understanding of potential GDP and the output gap.[2] Under the assumption that cyclical conditions are orthogonal to long-term growth (which is common in many economic models), this framework provides the right benchmark for sustainability. But if potential output is affected by cyclical conditions, what we call hysteresis, then the framework is not reliable and can produce excessively procyclical policy with damaging long-term consequences.

Fatas and Summers (2017) provides evidence of hysteresis during the fiscal contraction of 2010-2011 in Europe. Countries that implemented stronger consolidations saw a much larger reduction in potential output.[3] In the presence of hysteresis, not only we are underestimating the effects of fiscal policy on output, but we might fall in a vicious cycle. Low GDP growth is seen as structural and pushes policy makers to believe that further fiscal policy adjustments are needed. What is worse is that the successive rounds of contractions cause further reductions in potential output that validate the initial pessimistic and unfounded expectations of policy makers. In that world it is likely that we will end up with self-defeating fiscal contractions: while their goal was to improve sustainability, debt-to-GDP ratios are higher than before, because of the negative effects on GDP.

This paper uses the experience of European countries during the years 2008-2014 to understand the potential negative loop resulting from the use of estimates of

[2] The use of the output gap is central to the calculation of cyclically-adjusted fiscal variables at the national or European levels where fiscal variables are measured as a percentage of potential output. Estimates of potential output growth are part of the Debt Sustainability Analysis (DSA) of the IMF and similar assessments by the European Union.

[3] The evidence has been confirmed by Gechert et al. (2017).

potential GDP to steer fiscal policy when that fiscal is likely to generate negative effects on long-term GDP.

Section 2 reviews the goals that drive the design of fiscal policy and the role of potential output. Section 3 links that analysis to the EU fiscal policy framework. Section 4 analyzes the behavior of potential output estimates during the 2008-2017 period. Section 5 looks at how those estimates affected the design of fiscal policy and its potential consequences on GDP. Section 6 concludes and provides policy recommendations.

2. Fiscal Policy: Sustainability and Stabilization.

When setting fiscal policy plans, governments need to satisfy two goals. First, they need to ensure that budgetary plans are sustainable as measured by a non-explosive debt-to-GDP ratio. Second, they need to take into account cyclical conditions and provide the necessary stabilization policy. In both of these assessments, there is a need for a long-term forecast of GDP, which normally is expressed as an estimate of today's potential GDP. In the case of the sustainability assessment, long-term GDP growth rates feed into tax revenues but they also mechanically reduce the debt-to-GDP ratio. In the case of cyclical policy, potential output and the corresponding output gap provide an assessment of the cyclical state of the economy which is needed to design countercyclical budgets.

The medium-term objective

When it comes to sustainability, we typically define a steady-state level of the primary balance that ensures a sustainable path for government debt. Some notation: let D_t be the level of government debt, G_t spending, T_t taxes and Y_t the level of GDP in year t.

The budget balance (BB) can be written as

$$BB_t = T_t - G_t$$

And the primary balance (PB)

$$PB_t = BB_t + r_t D_t$$

Where r_t is the interest rate faced by the government.

For simplicity we will assume that our goal is to ensure that the debt-to-GDP ratio remains constant and equal to the level in year t. To design a sustainable policy we need to have an understanding of the long-term path of GDP and this is typically done by forecasting a growth rate for potential output Y_t^p (let's denote this steady-state growth rate by g).

We start with a situation where GDP is at potential ($Y_t = Y_t^p$), a zero output gap. Then, the government needs to maintain a primary surplus as a % of today's GDP equal to:

$$PB_t = (\mathrm{r} - \mathrm{g})\,\frac{D_t}{Y_t}$$

If we look at a case where taxes are proportional to GDP with a tax rate equal to τ and that the initial primary surplus matches the target one, the government can ensure sustainability by a spending rule that sets a path of government spending growing at a rate equal to the growth rate of potential GDP.

If the economy is not at potential, then the growth of GDP might be different from potential growth for a few years and this needs to be incorporated in our analysis.

In summary, in a sustainability exercise estimates of trend GDP matter for the steady-state growth rate as well as for the future temporary growth resulting from deviations of current level of GDP from potential.

Dealing with changing cyclical conditions

Under the assumption that fiscal policy is a useful stabilization tool, budget balances need to react to cyclical conditions. In the presence of a change in GDP, policy makers have to decide how much of this reduction in GDP is temporary versus permanent by estimating potential output and the corresponding output gap.

What is the right fiscal policy stance during cyclical events and how do we measure it? We first need to assess how fiscal policy affects economic activity. Measuring the stabilizing effect of fiscal policy requires us to have an economic model in mind. Blanchard (1993) suggests the use of inflation-adjusted budget

balance as a ratio to GDP as a proxy for the aggregate demand effect of fiscal policy in a given year.[4]

When looking at the change in the budget balance, it is common to produce cyclically-adjusted measures of the budget balance (CAB) to separate automatic stabilizers from discretionary actions. Let's assume for the moment that taxes and spending can be expressed as functions of output and that the CAB is simply expressed as the budget balance that would exist if output were equal to potential:

$$CAB_t = T\left(Y_t^p\right) - G\left(Y_t^p\right)$$

where $T(\cdot)$ and $G(\cdot)$ represent how revenues and spending depend on the level of economic activity.

The CAB is not observed so the calculation of cyclically adjusted balances requires an indirect approach. We start with the values of taxes and spending that can be observed and then assume a function of how taxes and government spending are automatically affected by the business cycle, as captured by deviations from potential output. Their cyclically adjusted counterparts are

$$T_t^{CAB} = T\left(Y_t^p\right) = T_t\left(\frac{Y_t^p}{Y_t}\right)^{\epsilon_T}$$

$$G_t^{CAB} = G\left(Y_t^p\right) = G_t\left(\frac{Y_t^p}{Y_t}\right)^{\epsilon_G}$$

[4] This is clearly the case in a simple static IS-LM model. Spending directly affects aggregate demand. Taxes can help stabilize disposable income and although the effect is not one-to-one because the marginal propensity to consume is lower than one, using the budget balance is close enough. Moving from a static model to a dynamic model, the relationship between fiscal policy and aggregate demand becomes more complicated. Now what matters for demand and output is not only current but also future fiscal policy. For example, to understand and measure the effects of a change in fiscal policy we need to assess how these changes translate into expected changes in spending and taxes, as well as how these affect other components of aggregate demand and potentially have an effect on the supply side of the economy. Blanchard (1993) shows that the previous result also applies to a simple intertemporal model that deviates from Ricardian equivalence, under the assumption of stable expectations regarding future taxes.

where ϵ_T and ϵ_G are the elasticities of taxes and spending relative to potential output.[5]

We can think of the automatic stabilizer part of the budget as the difference between the budget balance and the cyclically-adjusted balance measured as % of GDP (or potential GDP).

$$as_t = bb_t - cab_t = \frac{BB_t}{Y_t} - \frac{CAB_t}{Y^p} =$$

$$= \frac{T_t}{Y_t}\left[1 - \left(\frac{Y^p}{Y_t}\right)^{\epsilon_T - 1}\right] - \frac{G_t}{Y_t}\left[1 - \left(\frac{Y^p}{Y_t}\right)^{\epsilon_G - 1}\right]$$

From a stabilization point of view, the distinction between automatic stabilizers and discretionary policy is not relevant, what matters is the change in the budget balance. But their respective evolution matters for assessing whether sustainability is being put at risk or not. Automatic changes will correct by themselves when the output gap closes while changes to structural budgets require a future change in fiscal plans

Cyclical versus permanent shocks

Let's start with the budget balance at the medium-term objective level. Then, in the presence of a cyclical negative shock, a recession, the budget balance declines because of automatic stabilizers. If there is no discretionary policy, the structural balance will not change and we expect budget balances to return naturally to its medium-term target as the recessionary effects die out.[6] If, in addition, there is a countercyclical discretionary change in the budget balance then we will need to plan for a fiscal adjustment to reverse it, assuming we want to return to the same debt-to-GDP level. This adjustment does not need to take place at the same speed at the economy is returning to potential, but it needs to be consistent with the sustainability of current and future budgetary plans.

[5] For the sake of simplicity, we ignore interest payments from this equation. Hence, we do not discuss the differences between overall and primary balance.

[6] Although, in the presence of a negative shock, this cyclical decrease in the budget balance will increase debt and might trigger a need to readjust future budgetary plans, assuming that the previous level of debt is considered optimal. We will ignore this point in our following analysis.

In the presence of a pure permanent shock, there is no need for stabilization policy, the government should not let the automatic stabilizers work and should maintain the budget balance constant – which is also consistent, in this case, with a constant structural balance. How this gets implemented depends on how automatic stabilizers are defined. In some cases, they are a function of cyclical variables (such as unemployment), and they might not react to such a permanent change in output. In other cases, as taxes react to income, there will a decrease in the tax revenues that will require an adjustment in tax rates or spending to restore the desired primary balance.[7]

In most cases, there is a combination of a transitory and a permanent component. Either because shocks themselves have both components or because a pure supply (permanent) shock can generate some cyclical dynamics. In this case we need to separate the two components to know how much the fiscal balance should be allowed to move in order to provide the necessary stabilization function. And let's stress again that the evolution of the structural balance matters for setting future fiscal policy. A worsening of the structural balance means the need for future adjustment. In that sense the structural balance is both a target and an indicator for future policy. For both reasons, getting the estimates of potential output right is crucial.

The costs of mismeasuring potential output

Imagine the case of a pure temporary shock that is fully interpreted as a permanent shock. Authorities will strive to keeping both the budget balance and the structural balance constant but this means a wrong fiscal stance, as it will be procyclical. And the procyclicality will depend on our error measuring potential output. Had we measured the structural balance using the correct level of potential output, we would be seeing that there has been an increase in the balance, the signal of a procyclical policy. We can represent the procyclical fiscal stance (PFS) of such an error in measuring potential output using the formula above for automatic stabilizers

$$PFS_t = \frac{T_t}{Y_t}\left[1 - \left(\frac{Y^p}{Y_t}\right)^{\epsilon_T - 1}\right] - \frac{G_t}{Y_t}\left[1 - \left(\frac{Y^p}{Y_t}\right)^{\epsilon_G - 1}\right]$$

[7] In addition, a permanent shock to GDP also requires a future adjustment to structural balances (as the debt-to-GDP ratio is higher than planned).

Where Y^p is the correct value of potential output.

The procyclical nature of fiscal policy will have an effect on GDP growth via the standard multipliers so we expect output to fall by an amount

$$\Delta Y_t = \mu\,(-PFS_t)$$

Where μ represents the short-run multiplier.[8]

There is another potential source of procyclicality that is also related to the estimate of the output gap. As argued above, the structural balance is not just a target, it is also a central indicator of how sustainable fiscal policy is. If the elasticities used to calculate cyclically-adjusted variables vary with the cycle, we could also find that the mismeasurement of these elasticities generates procyclical fiscal policy. Imagine a temporary recession where, because of mismeasurement in the tax or spending elasticities, we wrongly conclude that the structural balance is declining, even if only automatic stabilizers are operating. This will also trigger a future consolidation even if it is not needed.

If shocks to GDP are symmetric and the mismeasurement of potential output or the output gap are also symmetric, all these dynamics will simply result in additional volatility in output but will not result in any change in growth or long-term GDP levels.

Potential output and fiscal consolidation in the presence of hysteresis

We now open the door for a much larger negative effect of mismeasuring potential GDP by allowing for the presence of hysteresis. While the term hysteresis was first used to describe the persistence of European unemployment during the 1970s (Blanchard and Summers (1986)), the notion of cyclical shocks leaving permanent effects is a much broader concept. When growth is endogenous and its driving process (investment, R&D) is possibly affected by cyclical conditions, then hysteresis is present across a variety of macroeconomic models (Stadler (1990), Fatás (2000)). Hysteresis should simply be interpreted as the years of growth that are permanently lost. The literature has recently

[8] And remember that the debt-to-GDP ratio will go up with the permanent shock and will also require a fiscal adjustment to restore previous level of debt. This consolidation effort will add to the procyclical stance of fiscal policy, generating an even more contractionary effect on GDP.

regained interest because of the persistence of GDP after the 2008 crisis (Rawdanowicz et al. (2014)).

When hysteresis is present, the concept of potential output becomes much harder to define as it is itself affected by cyclical conditions, including those related to fiscal policy decisions. Estimates of potential output as a long-term forecast of GDP cannot be made without a full understanding of economic policies decisions and their effects. This create a loop that makes the estimates of potential output a function of the fiscal policy which itself depends on those estimates of potential output.

Imagine a crisis is forecasted for next year so that GDP is expected to be below potential $Y_t < Y_t^p$. The government decides to engage in a fiscal consolidation because it misinterprets this GDP fall as a change in potential output.[9] The adjustment can take many forms but for simplicity let's assume that the fiscal adjustment takes the form of a one-year decrease in spending relative to its normal path:

$$FA_t = \Delta G_t = G_t - G_{t-1}(1+g) < 0$$

Where we use the notation Δ to refer to changes relative to previous planned levels and FA_t is the fiscal adjustment planned for t which we assume matches its execution.

The change in spending is likely to affect negatively GDP next year. The cyclical change in GDP (relative to previous forecasts) will depend on the fiscal policy multiplier (μ), as discussed earlier.

$$\Delta Y_t = \mu \, FA_t$$

Because of the consolidation efforts, the level of debt will be reduced, relative to its previous path, by

$$\Delta D_t = \Delta G_t - \Delta T_t = \Delta G_t - \mu\tau\Delta G_t = (1 - \mu\tau)\Delta G_t$$

And we know that $\Delta D_t < 0$ as long as $(1 - \mu\tau) > 0$.

[9] It can also be that the consolidation is triggered by the transitory decrease in GDP resulting in a temporary decrease in taxes and an increase in debt levels. It can also be that the consolidation is triggered by the reaction of financial markets to the crisis.

This reduction in the level of debt changes the burden on future government balances by

$$(\mathrm{r}-\mathrm{g})\,\Delta D_t = (\mathrm{r}-\mathrm{g})\,(1-\ \mu\tau)\,\Delta G_t$$

This reduction in the burden is the one that governments are hoping to get to bring budgetary plans back to a sustainable path

But so far we have assumed that potential GDP continues to grow at a rate g and that by time $t+1$ we have

$$Y_{t+1}^{p} = Y_t^{p}(1+g)$$

But what if potential output has been affected? What if its growth rate was transitorily lower during the crisis year?

Let us assume that some of the effects of the recession become permanent. In particular potential output changed by an amount (ΔY_t^p) that is related to the cyclical change in output (ΔY_t^c).

$$\Delta Y_t^{p} = \eta\ \Delta Y_t^{c} = \eta\ \mu\ \Delta G_t$$

Where the parameter η is the hysteresis parameter. Because of this change in output there will be a permanent loss of revenues equal to

$$\tau\ \Delta Y_t^{p} =\ \tau\ \eta\ \mu\ \Delta G_t$$

This means that at time $t+1$, the fiscal authorities will downgrade their estimates of potential output. In other words, they validate their wrong expectations. At the same time, they will realize that their fiscal adjustment was not enough to control the debt-to-GDP ratio. It is then likely that an additional adjustment will be required.

The presence of hysteresis changes the way we think about stabilization policy. The fact that potential output can be affected by cyclical developments means that stabilization policy should be much more aggressive (Blanchard, Cerutti, and Summers (2015)). Under this scenario, the goal of stabilization policy should be to run a "high pressure" economy in order to maximize long-term GDP. This is a strategy that is very different from the one suggested by symmetric models

of the business cycle without hysteresis where potential output is immune to business cycles and stabilization policy is simply about stabilizing GDP around a trend.

Self-defeating fiscal consolidations

If our arguments are right, not only fiscal consolidations can have long-lasting damage, but they can also lead become a self-defeating policy. Fiscal consolidations are designed to reduce debt and put fiscal policy on a sustainable path. But, from our previous discussion, this can have negative consequences on output and make the fiscal consolidation less successful as they will have an impact on tax revenues. Is it possible that these effects are so large that the fiscal consolidation is self-defeating?

DeLong and Summers (2012) discuss this possibility. Following our notation above, a fiscal contraction is self-defeating as long as the output costs impose a larger fiscal burden than the savings associated to the reduction in government spending. This means in practice, that

$$\tau\,\eta\,\mu \geq (\mathrm{r}-\mathrm{g})\,(1-\mu\tau)$$

We can express this intuition as an upper bound for the interest rate faced by governments as a function of all the other parameters. Fiscal consolidations are self-defeating as long as the interest rates is smaller than

$$\mathrm{r} \leq \mathrm{g} + \frac{\tau\,\eta\,\mu}{(1-\mu\tau)}$$

DeLong and Summers (2012) do a calibration for the US economy and conclude that the fiscal consolidation in the years that followed the 2008 crisis was very likely self-defeating. Fatas and Summers (2017) provide a similar calibration for the European economies for the 2010-11 fiscal consolidation to reach similar conclusions. The reason for such a pessimistic view is a combination of large fiscal multipliers (μ) and, more importantly, the estimated hysteresis parameter in the data (η). The degree of hysteresis estimated in Fatas and Summers (2017) is about 5 to 10 times larger than what is needed in the calibration.

Whether fiscal consolidations are self-defeating or not is not central to our argument. Our concern is on the endogeneity of potential output and how it leads to suboptimal fiscal policy. In the worst case scenario, if the reaction of

potential output is large enough, not only we see permanent scars of fiscal contractions but, in addition, the net effect of these consolidations is to raise the debt-to-GDP ratio.

3. Potential Output in the EU Fiscal Policy Framework

The EU fiscal policy framework was built around the principle of reducing the risks of governments following unsustainable fiscal policy plans. The original EU Stability and Growth pact, built on Articles 121 and 126 of the Treaty on the Functioning of the European Union provided a level of fiscal monitoring surveillance based on a simple set of numerical 3% and 60% caps for deficits and debt. Combined, they provide a framework for sustainability, under reasonable assumptions on growth. As time passed, the framework has moved away from its original simple rules and added flexibility and consideration for country-specific elements. Because of these changes the role of potential output grew over time.

In the 2005 reform of the Stability and Growth Pact the framework added a medium-term objective that could vary depending of the circumstances of the country such as initial debt but also potential growth. The combination of all these changes led to a much more systematic analysis of traditional debt sustainability.[10] As a result, it introduced the production of a regular Debt Sustainability Monitor report (twice a year) with a detailed government debt sustainability exercise.[11] These reports followed the standard textbook logic of debt sustainability and they are broadly consistent with the methods used by other organizations such as the IMF.[12]

It was the same 2005 reform that made explicitly the role of economic factors in justifying deviations from the 3% limit.[13] While the cyclically-adjusted balance had been used before as an additional indicator, the 2005 reform put it at the center of the analysis, as a tool to understand the role of special economic circumstances.

---

[10] See European Commission (2005)

[11] See European Commission (2014) for a detailed description of the methodology.

[12] The revised framework of the IMF was introduced in International Monetary Fund (2011).

[13] A 2002 European Commission Communication later adopted in 2003 had already made explicit the use of cyclically-adjusted balances as a complementary tool. See Turrini and Larch (2009)

The approach to measuring potential output and the output gap in the European Union has been driven by a production function approach. Each of the member countries also produce macroeconomic scenarios for their economies as part of their Stability and Convergence Program updates. These forecasts involve a combination of judgement and models similar to the EU common methodology (Mc Morrow, Roeger, and Vandermeulen (2017)). Similar production function models are used by the OECD. The IMF has traditionally relied more on judgement, partly because of the larger and more diverse group of countries over which forecasts need to be made (IEO (2014) but is recently moving towards a common centralized methodology, closer to the OECD and European Union (IMF (2015).

The EU fiscal policy framework also contains an alternative estimate of potential growth that is used in the expenditure benchmark. This alternative estimate is calculated using a 10-year window with estimates of the past 5 years combined with European Commission forecasts for the next 4. This series is designed to be smoother relative to the potential output used to calculate output gaps and structural balances.

The methodology used by the European Commission as well as by the other institutions is built on traditional macroeconomic models where the long-term dynamics are independent of the short-term dynamics. Short-term forecasts using judgement over the next 2 years are combined with the trends estimated by the production function. The trend produces a 5-year forecast and then the GDP forecasts over the next five years are required to reflect a closing of the output gap (output returns to trend).

Several assessments of these methodologies have reached the conclusion that potential output estimates react too much to cyclical changes in GDP. They tend to be overoptimistic during good years and over pessimistic during bad years (Mc Morrow, Roeger, and Vandermeulen (2017) or Kuusi (2017)). But in these assessments the cost of procyclical potential output estimates is just procyclical fiscal policy and more volatile GDP. As we have argued before, and in the presence of hysteresis, this procyclical policies can create long-term effects on GDP. In this case, the future evolution of GDP might wrongly validate the unfounded forecasts. In some sense, it is the economic policies designed by those forecasts the ones that are ensuring that they become accurate. In the next section we provide empirical evidence that all these ingredients are important to understand the conduct of fiscal policy in the EU during the 2008-14 period.

4. Shifting Goalposts During the Global Financial Crisis.

Since 2008 we have witnessed among advanced economies a period of low GDP growth that resulted in successive downward revisions to our GDP forecasts. And as the crisis developed, governments and international organizations slowly changed their long-term forecasts of GDP, as captured by potential output estimates. While this is not unique to this crisis, the succession of revisions to potential GDP was larger than usual. As shown early in Figure 1, relative to the pre-crisis trend, GDP in the Euro area remains about 15% lower.

In this section we explore how this revision to long-term performance happened. How long did it take for the crisis to change our long-term forecasts? Our ultimate goal is to understand how these worsening expectations of potential output affected fiscal policy plans in real time.

To study revisions to long-term GDP forecasts we make use of two sources.[14] We look at vintages of the IMF World Economic Outlook (WEO) during the years 2007 to 2018.[15] For each year we focus on the April issue of the WEO. Each of these vintages provides 5-year forecasts for GDP as well as potential GDP. We take the data from the year before as known and then we construct forecasts for up to 6 years later. For example, in April 2010, GDP for the year 2009 is taken as an actual number (even if it might be revised later) and we have forecasts all the way to 2015.

We also make use of the EU AMECO database. The AMECO database combined with the GDP forecasts of the European commission provides actual and potential GDP growth forecasts two years ahead. We use of data from 2007 to 2017 and always pick the Spring vintage, to be consistent with the IMF data. AMECO provides two separate long-term GDP forecasts: potential output and trend output. We look at both although the time series is shorter for trend output.

[14] Countries included in the two sources are listed in a data appendix.

[15] We start our analysis with the April 2007 issue because it precedes the decrease in growth rates we witnessed at the end of 2007.The NBER declared December 2007 as the starting month for the US recession. The CEPR concluded that the Euro had entered a recession in the first quarter of 2008.

We use the following notation for the value of a variable Y for the year $t+i$ that appears in the vintage of year t. If the value refers to a future year ($i > 0$) then this will be a forecast.

$$Y_{t+i}^{t}$$

For GDP in 2011, the forecast made in 2010 is expressed as

$$GDP_{2011}^{2010}$$

We now want to understand how accurate these forecasts were compared to actual data a few years later by calculating forecast errors. We look at one-year, two-year and six-year windows. As an example, take the 2010 vintage of AMECO or the WEO and look at a tw-year window. We take 2009 as known and we ask how the forecast for 2011 compares to the number reported for 2011 in the April 2012 vintage. In other words, we are comparing the forecast to the *closest vintage* when the data was already known. This gives us a perspective on how our views on the future evolved in *real time*. The 2-year forecast error is:

$$FE_{GDP,2011}^{2010,2012} = gdp_{2011}^{2012} - gdp_{2011}^{2010}$$

Where small letters represent logarithms. Notice that GDP_{2011}^{2012} is not a forecast, it is simply the GDP in 2011 according to the 2012 vintage of the database.

Because of data revisions, changes in base year and also changes in national accounting rules, the level of GDP might not be comparable across different vintages.[16] One way to avoid data revisions of GDP levels is to rewrite the expression for forecast errors in terms of GDP *growth*.[17] The forecast error for 2011 using the forecast made in 2010 then simply becomes the difference between actual and forecasted growth rates between 2009 and 2011

$$FE_{GDP,2009}^{2007} = (gdp_{2011}^{2012} - gdp_{2009}^{2012}) - (gdp_{2011}^{2010} - gdp_{2009}^{2010})$$

Calculating forecast errors for potential output is more complicated than for GDP. Potential output is not observed but estimated. In addition, revisions to current level of potential output tend to lead to revisions of ***past*** levels of

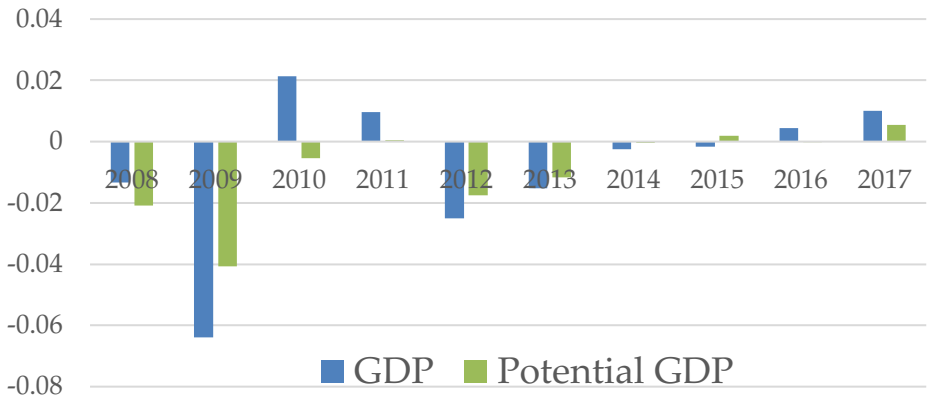
[16] For example, since October 2014 the WEO has started using updated data using ESA2010 criteria.

[17] This is the same approach followed by Blanchard and Leigh (2013). An appendix at the end of the paper describes in detail the calculation of forecast errors for actual and potential GDP.

potential output. In our calculations we ignore these historical revisions. What we are comparing is how our long-term forecasts of GDP change as time passes and for that we need a comparison of the actual *level* of potential output. Using the growth rates would underestimate the revisions to our forecasts. We explain in detail the methodology we use to deal with ex-post revisions of potential output estimates in an appendix at the end of the paper.

We present in Figure 2 the weighted average for all Euro members of the 2-year forecast error using the IMF WEO.[18] We include both GDP and potential GDP. Figure 2 clearly displays the economic shocks that characterized the 2008-2009 and 2012-2013 recessions. One thing to note is that we measure the *immediate* forecast error to GDP growth, not the revision relative to the first year in the sample. For example, the forecast error in 2012 corresponds to the forecast error for the period 2010-2012 when we compare the 2013 and 2011 vintages. If we were to measure the forecast error of the same year (2012) relative to the forecast in April 2008, the error would have been much larger. This is the reason why in 2010 and 2011 we see some positive surprises in GDP growth rates. We are simply revising upwards the pessimism that dominated 2008-2009.

Figure 2. Forecast Errors over 2-year horizon. Euro aggregate.



The second observation is that potential output reacts almost immediately to changes in GDP. Within our 2-year windows there is an almost perfect correlation between the forecast errors for the two variables. There is a small sense of asymmetry in the pre-2014 period as the positive surprises in 2010 and

[18] We use as weights the GDP in 2008. We maximize the number of countries included so not the same list of countries are included in the calculation of the Euro average for all years. For example, Estonia or Latvia are missing data on potential output for some of the early vintages of the WEO.

2011 did not lead to the same immediate revision. In the last years we do see more of that correlation even when surprises are positive.

To highlight this correlation between revisions to potential GDP and GDP surprises we make use of the individual countries in our sample and we calculate 1-year, 2-year and 6-year forecast errors for a sample of all advanced economies for which data is available in the WEO. We then do the same 1-year and 2-year calculations for the sample of European countries in AMECO.

We run panel regressions of forecast errors of potential output on forecast errors of GDP. Because potential GDP is a constructed variable, its forecast errors should be interpreted as how our views of long-term GDP changed in response to a surprise in GDP growth. When we look at 2-year and 6-year forecasts, we only pick non-overlapping years.[19] In all our regressions we include country fixed effects as well as year dummies.[20]

Table 1. Revisions to Potential GDP (IMF and AMECO)

	1-Year Forecast Error Potential GDP				
	IMF WEO			AMECO	
	Advanced	Europe	Euro	Europe	Euro
	(1)	(2)	(3)	(4)	(5)
1-Year Forecast Error GDP	0.648***	0.671***	0.735***	0.853***	0.859***
	(0.0688)	(0.114)	(0.0860)	(0.0591)	(0.0753)
Constant	-0.00504**	-0.00638**	-0.00494**	0.00639***	0.00653***
	(0.00201)	(0.00268)	(0.00191)	(0.00169)	(0.00243)
Observations	370	269	185	232	152
R-squared	0.555	0.534	0.657	0.704	0.722
Countries	35	26	18	30	19

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Our approach is very similar to the estimations of Coibion, Gorodnichenko, and Ulate (2017) or Blanchard, Cerutti, and Summers (2015) who also look at the persistence of cyclical events but using a more structured approach by identifying specific shocks.

[19] Although the overlap in information is not perfect given that, for example, the forecast for growth of GDP in 2008 done in April 2007 and in April 2008 contain different information.
[20] Results are practically identical if we remove year dummies or fixed effects.

Table 1 shows the result for three separate samples: Advanced, Europe and the Euro area.[21] There is strong consistency across the five columns. there is a very strong relationship between the two variables and the coefficient is around 0.65-0.86, signaling that surprises in GDP lead to large immediate revisions to our long-term views on output.

We now explore whether these revisions to potential output are asymmetric across positive and negative shocks. We separate forecast errors on GDP between those that are positive and those that are negative and columns (1) to (5) of Table 2 replicate some of the results of Table 1.[22]

Table 2. Revisions to Potential GDP. Asymmetric?

1-Year Forecast Error Potential GDP					
	IMF WEO			AMECO	
	Europe	Euro	Euro pre-2014	Europe	Euro
	(1)	(2)	(3)	(4)	(5)
1-Year Forecast Error GDP+	0.496**	0.553***	0.173	0.837***	0.851***
	(0.180)	(0.104)	(0.276)	(0.0703)	(0.0676)
1-Year Forecast Error GDP-	0.890***	0.914***	1.015***	0.961***	1.021***
	(0.130)	(0.142)	(0.214)	(0.0923)	(0.120)
Constant	-0.00502**	-0.00386*	-0.00136	0.00619***	0.00640**
	(0.00218)	(0.00216)	(0.00247)	(0.00176)	(0.00240)
Observations	269	185	113	232	152
R-squared	0.544	0.664	0.625	0.706	0.725
Countries	26	18	18	30	19

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

There is a sense of asymmetry in the revisions to potential GDP in the period 2007-2017. Positive changes to GDP are not translated into an increase in

[21] In the regressions we try to maximize the number of observations (countries) even if the panel is not always balanced. The number of countries for which some data is missing is small and removing them from the estimations do not change any of our results. The Europe or Euro area samples do not exactly match across the IMF and AMECO databases. The appendix lists the countries in each of these samples.

[22] We focus on the Europe and Euro samples (results for the group of advanced economies are very similar).

potential GDP as much as negative ones. The effect is strong in the IMF database but less so in the AMECO databases (although similar calculations with a 2-year forecast window reveal larger asymmetries in the AMECO database). The effects are even more pronounced during the 2007-2014 subsample. This is, of course, a special period of time where most GDP forecast errors are negative, but it gives a useful perspective on the continuous pessimism that was translated during these years into more negative views of long-term forecasts of GDP.

We have also estimated the relationship between GDP forecast errors and revisions to potential output by using a longer window. We choose a 2-year window that corresponds to the longest possible window using the AMECO database. We also look at the longest possible window for the IMF database (6-years). As an example, the April 2007 WEO vintage gives us forecasts up to 2012. We then calculate the forecast error for the period 2006-2012 for both GDP and potential GDP. We do that as well for the April 2012 vintage that includes forecasts for 2017 (there is a one-year overlap in the two windows but that's the only way we can get two six-year forecast errors). We then run a similar panel with fixed effects and time dummies as in the previous tables. Results are presented in Table 3.

Table 3. Longer horizon forecast errors for Potential GDP

Forecast Error Potential GDP (2 and 6-year)					
	IMF WEO			AMECO	
	Europe 2-year	Euro 2-year	Euro 6-year	Europe 2-year	Euro 2-year
	(1)	(2)	(3)	(4)	(5)
Forecast Error GDP	0.823***	0.877***	0.945***	0.812***	0.856***
	(0.0755)	(0.0539)	(0.0240)	(0.0859)	(0.103)
Constant	0.00797	0.00597	0.0163***	-0.00855***	-0.00857**
	(0.00535)	(0.00505)	(0.00347)	(0.00226)	(0.00314)
Observations	123	85	31	86	57
R-squared	0.868	0.912	0.995	0.706	0.760
Countries	26	18	17	30	19

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Overall, we confirm our previous results, but the coefficient is now slightly higher (more so for the IMF WEO). This would be expected as changes in GDP

that are more persistent should have a stronger influence on our views regarding potential output. But, at the same time the coefficient on the 2-year or the 6-yerar coefficient is not too different from the 1-year one, confirming that potential output estimates react very quickly to surprises in GDP. The results are almost identical for the WEO and AMECO databases.

We provide an additional robustness check in Table 4, column (1), by using trend GDP from AMECO, an alternative measure of long-term GDP that is used in calculations of the sustainable expenditure path in the fiscal surveillance framework. This measure is designed to be smoother than potential GDP. The coefficient is almost identical to the one from previous table (0.8), confirming that even the smoother measure of potential GDP strongly reacts to GDP surprises.

Table 4. Alternative measures of Potential GDP

2-Year Forecast Error Potential GDP		
	AMECO (Euro)	
	Trend GDP	Uncorrected Potential
	(1)	(2)
Forecast Error GDP	0.799***	0.356***
	(0.204)	(0.0785)
Constant	0.000994	-0.00828**
	(0.00310)	(0.00325)
Observations	38	57
R-squared	0.649	0.330
Countries	19	19

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In column (2) we present the results of using what we call the uncorrected measure of potential GDP. This is using the forecast error of the growth of potential GDP as measured directly from the AMECO vintages. But this forecast error is misleading because as time passes AMECO revised backwards its estimates of potential output for previous years. What this means is that the pessimism about the future also changed the past and generates an ex-post smoother series for potential. This has the effect of lowering "*ex-post forecasted* growth rates" and, as a consequence, reduces the forecast errors and the coefficients in the regressions. These results are not relevant to our arguments,

because what matters for fiscal policy is how the current *level* of potential GDP informs fiscal authorities about long-term GDP forecasts. Our previous regressions that had corrected for those ex-post revisions of history are the right way to measure this concept.[23]

We have now confirmed in all the previous tables that policy makers revised their estimates of potential output immediately after a GDP growth surprise. How accurate were these real-time views? Because potential output estimates are an indication of long-term GDP trends, we can check the accuracy of the policy makers views by comparing the 2-year forecast errors with the 6-year forecast errors from the IMF database. Do we see future surprises that represent a reversion to trend? Or does the persistence continue during the following years?

Table 5 shows the result using the April 2010 vintage. The two-year forecast errors get propagated over time and result in even larger forecast errors for later years. The fact that the coefficient is larger than one not only confirms that the revisions in potential output (April 2010) seemed to be warranted but it also hints that something else happened during the four years that followed that amplified the estimated persistence. In some sense, their pessimism turned out to be too optimistic. We provide a potential explanation for this pattern in the next section.

Table 5. Correlation forecast errors

6-Year Forecast Error Potential GDP (April 2010 WEO)		
	Europe	Euro
	(1)	(2)
2-Year Forecast Error	1.621***	1.848***
GDP (April 2010 WEO)	(0.418)	(0.413)
Constant	-0.0585***	-0.0648***
	(0.0137)	(0.0179)
Observations	26	18
R-squared	0.341	0.384

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[23] The appendix provides details on why our corrected potential output measure is equivalent to level comparisons of long-term GDP.

5. EU Fiscal Policy During 2007-2014 and Revisions to Potential Output

Fiscal consolidation in 2010-2014

How did surprises in GDP growth affected fiscal policy in the years 2007-14? As growth declined in 2008-09 budget deficits increased because of the functioning of automatic stabilizers. In addition, some governments engaged in discretionary fiscal policy and implemented stimulus packages, leading to a reduction in structural budget balances. Increasing debt-to-GDP ratios combined with larger structural deficits put countries in a trajectory that required an adjustment to current and future fiscal plans to bring fiscal policy back to a sustainable path.

Figure 3. Planned Change in Structural Balances % of GDP (Euro Area)

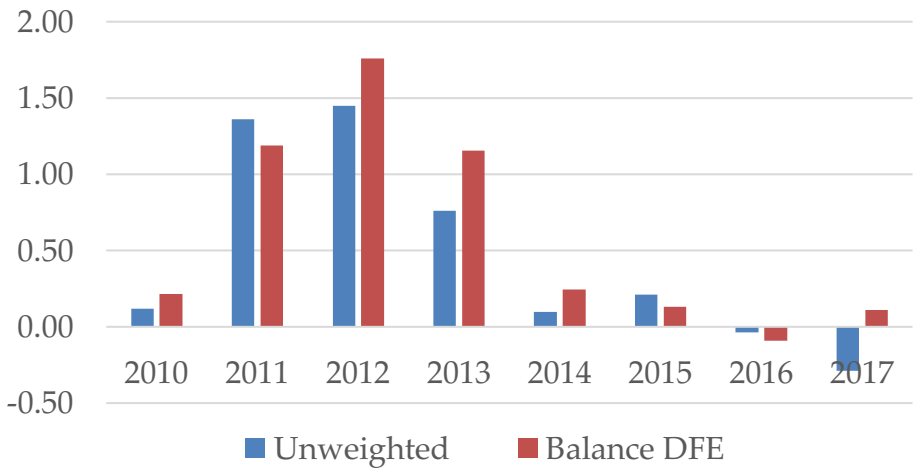
In Figure 3 we show how fiscal plans evolved during these years. We show 1-year forecasts of changes in structural balances (measured as % of GDP) for the average of Euro members. We calculate the change both as an unweighted average and one where countries are weighted by their GDP size in 2008. The data shows the planned change in structural balances during that particular year as reported in the April IMF WEO issue. So, for example, 2011 is the change in structural balance from 2010 to 2011 as reported in April 2011.

The figure presents the well-known fact that after the 2008-09 crisis, Euro countries engaged in a fiscal consolidation during the years 2010-2013. While

these figures correspond to the plans in April of that year, their evolution matches well the implemented changes. In particular, we can compare these changes to the "narrative" estimates produced by the European Commission as captured by the Discretionary Fiscal Effort (DFE) indicators.[24] The DFE indicator is the average for the Euro area as reported in the AMECO database while the planned change in structural balance is our unweighted average of changes reported for that year by the IMF WEO.

Figure 4 compares the two magnitudes and it is clear that the overall numbers are very similar even if there are differences because of their different methodologies (Carnot and de Castro Fernández (2015)). We see a wave of fiscal consolidation that starts in 2010 and continues until 2013. We will now study these years as two separate waves: 2010-11 and 2012-13.

Figure 4. Discretionary Fiscal Policy. Planned Structural Balance Change versus Actual DFE



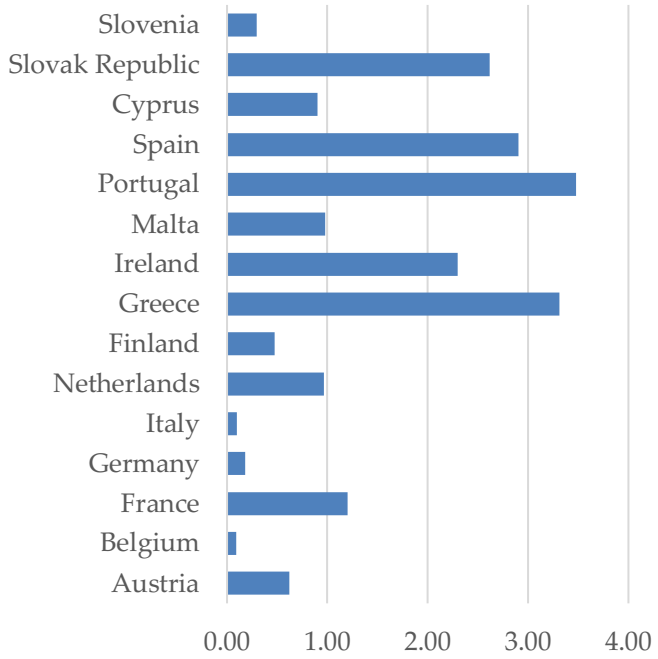
The first wave of fiscal consolidation (2010-2011)

The process of fiscal consolidation in the years following 2010 has received much attention in the academic literature because of the sudden and significant change in fiscal policy. This represented an opportunity to study the effects of contractionary fiscal policies by making use of the differences in the degree of

[24] The Discretionary Fiscal Effort indicators combine follow a bottom-up, narrative approach on the revenue side and a more standard top-down approach of the expenditure side. See Carnot and de Castro Fernández (2015).

fiscal consolidation across countries. There was indeed large variation in the data. For example, in 2011, while some countries had planned for a large fiscal consolidation (Greece, Portugal, Spain), others were looking at an almost constant structural balance (Germany, Belgium or Italy), as displayed in Figure 5.

Figure 5. Planned fiscal consolidation in 2011 (IMF WEO April 2010).



The work of Blanchard and Leigh (2013) is central to the literature on understanding the growth effects of the 2010-2011 fiscal consolidation. By comparing how planned fiscal consolidations during those two years correlate with changes in forecast errors of GDP allowed them to estimate the true fiscal policy multiplier.

We start by replicating their work. We use the same years, 2010 and 2011, where fiscal consolidations were planned and executed among many economies (as we can see from Figure 3 the consolidation in 2011 was much more significant than in 2010). We calculate the planned change in the structural balance as a percentage of GDP ($\Delta SB^{2010}_{i,2010-2011}$) as an indicator of the fiscal consolidation. We then regress the forecast error for real GDP for those two years on this measure.[25]

[25] We calculate the forecast errors relative to the GDP growth measured from the April 2012 WEO. This is slightly different from the way it is done in Blanchard and Leigh (2013) who used

$$FE_{i,GDP}^{2012,2010} = \alpha + \beta\, \Delta SB_{i,2010-2011}^{2010} + \varepsilon_i$$

Under the assumption that the forecast had been made using the right fiscal policy multipliers, the coefficient β should be equal to zero (i.e. the IMF was right in assuming a 0.5 multiplier). Blanchard and Leigh (2013) found that the coefficient was negative, large and significant, a sign that fiscal policy multipliers had been underestimated by the IMF model.

Table 6. Output growth effects of the 2010-11 fiscal consolidation

	$FE_{GDP}^{2012,2010}$	
	Europe	Euro
	(1)	(2)
$\Delta SB_{2010-2011}^{2010}$	-1.007***	-1.126***
	(0.229)	(0.217)
Constant	0.00876**	0.00805*
	(0.00402)	(0.00428)
Observations	23	15
R-squared	0.480	0.674

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The results of Table 6 are consistent with the view that fiscal policy multipliers were around 1.5-1.6, as established in previous papers in the literature.

We now want to understand how this change in GDP triggered by the fiscal consolidation translated into revisions to long-term GDP forecasts. In Table 7, first two columns, we regress the forecast error of potential GDP for the same years 2010-2011 also against the planned change in the structural balance from the April 2010 WEO. We see an effect that is large and significant. In many ways it confirms our previous results. A 1% change in the structural balance triggered a change in GDP of about 1.5% and translated into a change in estimates of potential output of about 0.8, which is about two thirds of the change in GDP and consistent with the estimates of Table 1. In other words, we learn that fiscal-policy induced changes in GDP are no different from any other change in GDP and affect in real-time estimates of potential GDP by a similar amount.

the latest available WEO to do that comparison. The reason for choosing 2012 is that later we plan to look at how the forecast errors in real time changed policy makers decisions.

To check the robustness of this result, in the last two columns we run a 2SLS specification where in the first stage we use the results of Table 6 as the fitted variable for this second regression. We can interpret this regression as first isolating the changes in GDP that were caused by fiscal consolidation and then in the second stage seeing how those changes affected our views on potential output. The coefficient is not far from our results of Table 1 where we had not isolated any particular shock. This methodology, also used in Fatas and Summers (2017) confirms that fiscal consolidation had an immediate effect on our long-term views on output.[26]

Table 7. Potential output growth effects of the 2010-11 fiscal consolidation

	$FE_{POT}^{2012,2010}$			
	Europe	Euro	Europe	Euro
	(1)	(2)	(3)	(4)
$\Delta SB_{2010-2011}^{2010}$	-0.732***	-0.821***		
	(0.192)	(0.246)		
$\widehat{FE}_{GDP,2011}^{2010}$			0.727***	0.729***
			(0.132)	(0.120)
Constant	-0.000867	0.000397	-0.00723***	-0.00547*
	(0.00337)	(0.00484)	(0.00231)	(0.00255)
Observations	23	15	23	15
R-squared	0.410	0.462	0.718	0.837

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This deteriorating views on GDP and potential are likely to influence fiscal policy in the years ahead. But there is also a related dimension that might affect fiscal policy: the behavior of inflation. This is normally ignored when looking at fiscal multipliers as we are interested in the reaction of *real* economic activity. But inflation matters when we are trying to understand the sustainability of fiscal policy. A surprised change in inflation will have an effect on the real value of the debt (which is denominated in nominal terms).

[26] In Fatas and Summers (2017) as well as Gechert et al. (2017) there is a variety of robustness tests over different horizons that confirm this result.

In Table 8 we replicate the results of Table 6 but instead of using the forecast error of real GDP we use the forecast error of *nominal* GDP. The fact that the coefficient is even larger suggests that prices move in the same direction as real activity (in the cross-section of countries). This means that fiscal sustainability deteriorated as a result of both the decline in real activity (which was perceived as persistent or permanent) and the decline in prices.[27]

Table 8. Nominal output growth effects of the 2010-11 fiscal consolidation

	$FE_{NGDP}^{2012,2010}$	
	Europe	Euro
$\Delta SB_{2010-2011}^{F,2010}$	-1.445***	-1.547***
	(0.237)	(0.238)
Constant	0.00761	0.00983
	(0.00497)	(0.00560)
Observations	23	15
R-squared	0.546	0.678

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

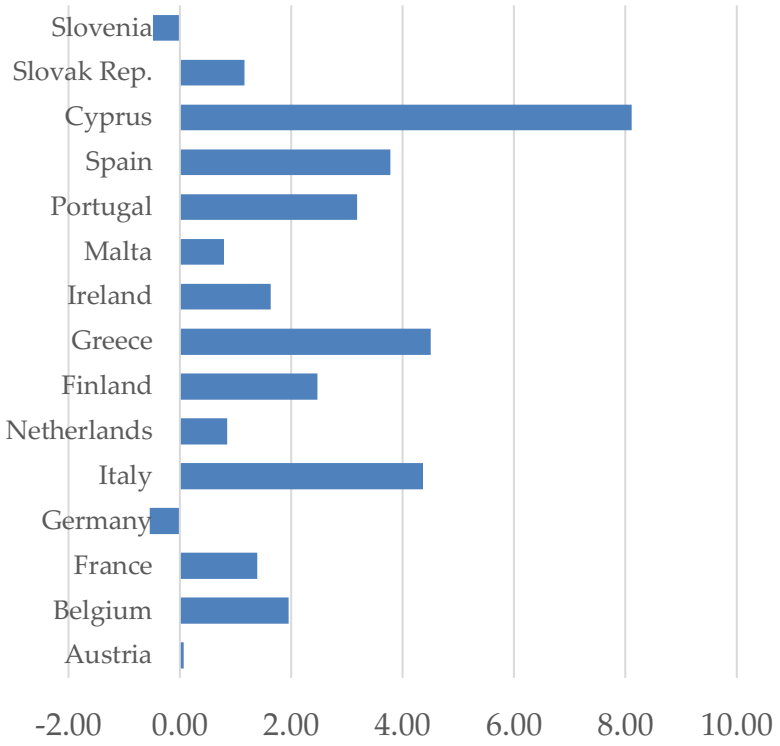
From the first to the second wave of fiscal consolidation (2012-2013)

We now ask how the pessimism of the 2010-11 years triggered by GDP surprises and revisions to estimates of potential led to changes in fiscal policy over the two years that followed. To capture this effect, we analyze changes in *future* consolidation plans across two vintages. In particular, we start with the April 2010 vintage and measure the plans for structural balance changes for the years 2012 and 2013 (these are the two years after the first wave of fiscal consolidation). We think about an increase in the balances between those years as the planned future fiscal consolidation. We then compare these plans with the ones that appear in the April 2012 vintage. How did our plans for those years changed between April 2010 and April 2012?

Figure 6 shows the results of that calculation. With the exception of Slovenia and Germany, we moved fiscal policy to a tighter stance. While we are looking at changes in the structural balance during those 2 years, the same results hold if we simply look at the planned level for the structural balance in 2013.

[27] These price effects were persistent over time. We have checked the persistence of these effects by replicating these regressions using our 6-year forecast errors.

Figure 6. Change in 2013-2011 Fiscal Consolidation Plans.
(Vintage 2012 vs Vintage 2010)



The fact that these plans changed between 2010 and 2012 should not be a surprise. It reflects the changing economic conditions in between the two vintages that required an adjustment in fiscal balances to restore sustainability. GDP growth was surprisingly low, and this changed our views on potential GDP. These changes were likely to trigger a tightening of fiscal policy over the following two years as we can see in Figure 6.

In addition, and from a stabilization point of view, as potential output is revised downwards, we develop a benign view of cyclical conditions. As a result, the tightening of fiscal policy does not seem as contractionary or pro-cyclical. This is the logic used by the European Commission and the Council to support the additional fiscal tightening in those years (see Claeys, Darvas, and Leandro (2016)).

We now present a quick test of whether this tightening was related to previous GDP surprises in the cross section of countries (Table 9). We regress the change in future fiscal consolidation plans against the forecast error for GDP for the

previous two years. Columns (1) and (2) focus on the revision of fiscal consolidation plans for the years 2011-2013 measured as the expected change in the structural budget balance (as % of GDP). And it compares the 2012 with the 2010 vintages of the IMF WEO. We regress that revision on the forecast error for GDP growth in the years 2011-2009, also across the two vintages. We are capturing how surprises in economic performance affected future fiscal plans.

The negative coefficient suggests that worsening economic conditions led to tighter fiscal policy (larger adjustments in structural balances), as expected. Interestingly, when we run the same regression for the whole period (including fixed and time effects) the coefficient is much smaller and insignificant. This reflects the special circumstances that surrounded those years where the pessimism, combined with other potential effects, pushed governments to a much tighter response to changing cyclical conditions.

Table 9. Changing Economic Conditions and Future Fiscal Plans

	Europe 2011-2013	Euro 2011-2013	Euro Full Sample
	(1)	(2)	(3)
Forecast Error	-0.230**	-0.270**	-0.0224
GDP (previous 2 years)	(0.109)	(0.117)	(0.0317)
Constant	0.0143***	0.0225***	0.00223***
	(0.00505)	(0.00582)	(0.000227)
Observations	23	15	146
R-squared	0.059	0.099	0.364
Number of countries	23	15	18

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9 has confirmed that worsening economic conditions led to a second round of fiscal tightening in the years 2012-13. Did this second round have the same negative effects on GDP and on estimates of potential GDP as the first one?

We first make use again of the methodology of Blanchard and Leigh (2013) in Table 10, which uses the same specification of Table 6 but for the period 2012-13. The results confirm the negative effects of that second fiscal consolidation. The size of the coefficient is smaller, but the implied multiplier remains higher than one.

Table 10. Output growth effects of the 2012-13 fiscal consolidation

	$FE_{GDP}^{2014,2012}$	
VARIABLES	Europe	Euro
	(1)	(2)
$\Delta SB_{2012-2013}^{2012}$	-0.638***	-0.641**
	(0.218)	(0.275)
Constant	-0.00952*	-0.0115
	(0.00540)	(0.00773)
Observations	24	16
R-squared	0.281	0.280

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

And for the sake of completeness we also replicate the result of Table 7 for this second wave of consolidation. Here we are looking at the effects of these contractionary policies on potential output estimates.

Table 11. Potential output growth effects of the 2012-13 fiscal consolidation

	$FE_{POT}^{2014,2012}$			
	Europe	Euro	Europe	Euro
	(1)	(2)	(3)	(4)
$\Delta SB_{2012-2013}^{2012}$	-0.642***	-0.654**		
	(0.217)	(0.273)		
$\widehat{FE}_{GDP,2013}^{2012}$			1.007***	1.020***
			(0.223)	(0.298)
Constant	-0.00465	-0.00552	0.00494	0.00622
	(0.00537)	(0.00767)	(0.00527)	(0.00829)
Observations	24	16	24	16
R-squared	0.285	0.291	0.692	0.651

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The first two columns show that the planned fiscal consolidation had a negative effect on our revisions to potential output forecasts that is almost the same size as the GDP forecast error, suggesting once again that cyclical conditions were perceived as leaving permanent scars in the economy. The last two columns use the 2SLS procedure that isolates movements in GDP caused by the fiscal consolidation (from Table 10) to identify how changes in GDP translated into

changes in potential output estimates. The coefficient is close to one suggesting that this translation was as strong or stronger than during 2010-11.

Our analysis of the period 2010-14 has revealed that the dynamics between fiscal policy and economic conditions led to a succession of fiscal tightening decisions that likely transformed unfounded pessimistic expectations into a reality of declining potential output growth.[28] The first phase of the consolidation during the years 2010-11 had clear negative consequences on our long-term views on GDP. This led to a second wave of consolidations that further reduced output and through hysteresis effects also potential output. These actions are likely to be irreversible as lower potential output becomes a measure of the foregone years of growth.

Finally, and as argued before, the hysteresis effects implied by these estimations are large enough that they are likely to be self-defeating.[29] Fiscal policy was designed to make budgetary plans sustainable but given the large long-term effects, it is very likely that we ended up with higher debt-to-GDP ratios despite the decrease in spending or increase in taxes. If this is the case, we are not talking here about a tradeoff between two desirable outcomes, we are talking about a clearly suboptimal policy.

6. Conclusions and policy recommendations.

The loop between fiscal policy and potential output

Our results present an analysis of the potential negative loop that is created by the interaction between overly pessimistic views on potential output and fiscal policy during recessions; a story that fits well the stylized facts of the Euro crisis in the years 2008-2014. Contractionary fiscal policy was designed using a long-term GDP forecast that only turned out to be right partly because of the negative effects of those same contractionary policies.

[28] Kuang and Mitra (2018) reach a similar conclusion using different methodologies.

[29] See Fatas and Summers (2017) for a map between hysteresis estimates and self-defeating fiscal consolidations. Auerbach and Gorodnichenko (2017) reach a similar conclusions using a larger sample of fiscal shocks among OECD economies. House, Tesar, and Pröbsting (2017) also confirm the persistent effects of fiscal policy changes in European countries and the possibility of self-defeating fiscal consolidations.

Any fiscal framework requires estimates of potential output and it is well known that any measurement errors will lead to the wrong fiscal policy, as acknowledged many times by European fiscal authorities. Where there is less clarity is on the consequences and importance of the mismeasurement.[30]

We first need to understand that there are two separate issues when talking about biased estimates of potential: whether *on average* forecasts are too optimistic and whether forecasts are too reactive to cyclical conditions and therefore generate procyclical policy. On the first issue this paper has nothing to add except that, of course, we want a methodology that is unbiased when forecasting long-term potential growth. The focus of our analysis has been on the second issue, the tendency of short-term growth fluctuations to influence long-term GDP forecasts and how it leads to a procyclical bias in fiscal policy.

For those who worry about fiscal policy procyclicality in good times, they see the optimistic bias as a risk to sustainability of debt (see Mc Morrow, Roeger, and Vandermeulen (2017), from the European Commission). Our focus was the opposite, as we zoomed in a recessionary episode where the procyclical bias had the effect of making the recession deeper and, even worse, affecting potential output and leaving permanent scars on the economy via hysteresis. And those permanent scars had the pernicious effect of validating unfounded pessimistic forecasts.

One could argue that this hysteresis effect is symmetric and during booms we could see the opposite effects. In that case, the actual consequences on the economy would be to add volatility without affecting its long-term path.[31] But this assumes a level of symmetry that is not supported by the data. First, business cycles are asymmetric. One extreme view argues that business cycles are just downward deviations from potential output as in Friedman's "plucking model" (Friedman (1993), Fatás and Mihov (2013)). Or we can also simply think about recessions being deeper than expansions (Neftçi (1984).

In addition, there is a second type of asymmetry that matters, related to the effectiveness of fiscal policy in booms and recessions. If fiscal policy multipliers

[30] As an example, in a pre-crisis analysis, Turrini and Larch (2009) acknowledge the weaknesses of the current methodology to measure the output gap but dismiss its importance and conclude that the issues can easily be addressed with minor tweaks to the methodology.

[31] This type of volatility would be very persistent but as long as symmetry holds, there will be no clear ex-ante bias in terms of whether this helps or hurts long-term GDP.

vary over the business cycle then procyclical fiscal policy will have small effects during expansions but could be quite damaging during recessions (Auerbach and Gorodnichenko (2011) or Jordà and Taylor (2016)). This asymmetry means that the potential hysteresis effects would be larger in the presence of negative shocks than in the presence of positive ones, having an overall negative effect on long-term GDP.

Our paper does not provide any evidence in favor of these asymmetries as we study one particular (negative) event. But our results support the view that the costs of procyclical fiscal policy during downturns are very large because of their effects on potential output. Given our estimates, we confirm and extend the interpretation of Fatas and Summers (2017), that this particular episode during the years 2010-2014 is likely to represent a case of a self-defeating fiscal consolidation because of the large long-term damage to the path of GDP.

We need to admit that this is, of course, a unique event because it was very large recession that included a financial and banking crisis and where the central bank was constrained by the zero-lower bound (and potentially had a too optimistic view on growth in 2011). So maybe one could argue that such an episode would not be repeated in the future and that we should not design a fiscal policy framework around it. We disagree. A fiscal framework needs to be designed to account for these large events given their significant consequences. These are the times when currency areas are tested. In addition, while in smaller recessions some of the effects we are highlighting here might not be as large, they will still be there and also need to be addressed.

Towards a better fiscal policy framework

The problem highlighted in this paper is one of procyclicality of measures of potential output. While this seems like a technical issue that should be improved through better modeling and econometric techniques, there is no yet a solution that appears ready to replace the current methodologies. We discuss here some recent proposals and the ways in which they address the weaknesses identified in this paper.

There is a first debate around which indicator to use to judge the appropriateness of fiscal policy. Today the indicator is the structural budget balance. The budget balance is the key variable to judge sustainability and it is also an indicator of the fiscal policy stance. The structural budget balance makes use of the same long-term GDP forecast as any estimate of potential or trend GDP. And, in principle,

any fiscal policy variable that relies on these estimates of potential output should share the same weaknesses. However, not every fiscal policy variable relates to potential output estimates in the same way. In particular, the structural balance can be a much more volatile variable because it also requires accurate calculations of elasticities of taxes and spending that have proven to be noisy and, in some cases, adding to the procyclicality of fiscal policy. Also, one-time changes in taxes or spending can easily distort the measure of the government balance.

As a result, many of the recent proposals for reform emphasize the advantages of expenditure rules combined with a debt-to-GDP ratio as an anchor. For example, Andrle et al. (2015) conclude that "the lowest variability of output is achieved by the expenditure growth rule, which ties down real expenditure growth to the economy's potential or trend growth rate, combined with a debt correction mechanism".[32] This logic is also supported by the proposals of Bénassy-Quéré et al. (2018) or Claeys, Darvas, and Leandro (2016).

But replacing structural balance rules with expenditure rules will not be a complete solution. One still needs an understanding of potential growth to enforce an expenditure rule.

An obvious reaction is that we need a smoother series of potential output. As Claeys, Darvas, and Leandro (2016) put it: it is hard to rely on a measure of potential output that gets revised annually by an amount that is larger than 0.5% of GDP, the required baseline annual adjustment within the EU fiscal policy framework. Andrle et al. (2015) present simulations where long-term GDP growth estimated as a moving average using the past 5 years plus forecasts for the next 4 provides additional stability to fiscal policy.[33] Although, as highlighted in Irish Fiscal Advisory Council (2015), an expenditure rule that relies on a such an estimate of trend GDP growth can also generate suboptimal policy at times when the trend growth series is volatile. One potential solution is to avoid

[32] Similar conclusions are reached by Kuusi (2017).

[33] There are some dissenting voices to the view that longer horizons are superior. For example, Mc Morrow, Roeger, and Vandermeulen (2017) are concerned that looking at forecasts over longer periods might just translate into even more optimistic forecasts during booms that will lead to more procyclicality. However, their concern is about procyclical behavior during booms and the fact that extending the number of years it is allowed to return to trend will lead to more procyclicality. Their argument does not apply in the presence of negative shocks. We do want to increase the number of years to avoid procyclicality.

putting too much weight on recent data to avoid an overaction to cyclical events (Kuang and Mitra (2018)).

But smoother estimates of potential output do not fully address all the concerns raised by our results. What we need is a better understanding of the nature and persistence of different shocks. We can potentially rely on economic models to inform us about this persistence and produce better real-time estimates of potential GDP. For example, Coibion, Gorodnichenko, and Ulate (2018) present several alternatives to this type of methodology that create very different results from the currently accepted methodologies. While promising, in order to implement any of their methods we would first need to reach consensus on the way different shocks and their effects are identified.

When it comes to this identification, there are two concerns raised by the results of this paper. First, how to handle positive and negative shocks. We have argued that there is an inherent asymmetry when it comes to the consequences of procyclical fiscal policy. Avoiding procyclicality during a recession is much more important than during an expansion. Having a rule that takes into account this asymmetry, in particular for large shocks, is not easy. It is likely that a certain amount of flexibility and judgement would be required for a proper decision, which bring us back to the debate of strict and simple numerical rules versus the use of judgement. While in other circumstances simple and numerical rules might dominate, this seems to be the case where flexibility via a clause for fiscal policy during "exceptional circumstances" should apply. Such a clause is already present in the EU fiscal policy framework but in order to be effective the potential costs associated to excessively tight fiscal policy during deep recessions needs to be better understood and accepted by policy makers.

And this brings as to the second issue, which is possibly more fundamental and harder to manage. We need to have a common understanding on the extent of hysteresis effects and how countercyclical policy should be used to minimize them. Standard macroeconomic models do not even contemplate this discussion. We hope that the results of this paper and other recent ones in the literature also showing strong support for the existence of hysteresis effects can helps us design new models as a step to design better fiscal policy. Recognizing the existence of hysteresis should make economic policies (fiscal and monetary) much more aggressive in particular during large negative cyclical events like the one the Euro area experienced during the 2008-2014 period.

7. References.

Andrle, Michal, Mr John C Bluedorn, Luc Eyraud, Mr Tidiane Kinda, Petya Koeva Brooks, Mr Gerd Schwartz, and Anke Weber, 2015, *Reforming Fiscal Governance in the European Union*. 15–19 (International Monetary Fund).

Auerbach, Alan J., and Yuriy Gorodnichenko, 2011, Fiscal Multipliers in Recession and Expansion. Working Paper, National Bureau of Economic Research.

Auerbach, Alan J, and Yuriy Gorodnichenko, 2017, Fiscal stimulus and fiscal sustainability, National Bureau of Economic Research.

Bénassy-Quéré, Agnès, Markus Brunnermeier, Henrik Enderlein, Emmanuel Farhi, Marcel Fratzscher, Clemens Fuest, Pierre-Olivier Gourinchas, Philippe Martin, Jean Pisani-Ferry, Hélène Rey, Isabel Schnabel, Nicolas Veron, Beatrice Weder di Mauro, and Jeromin Zettelmeyer, 2018, Reconciling risk sharing with market discipline: A constructive approach to euro area reform. CEPR Policy Insight.

Blanchard, Olivier, Eugenio Cerutti, and Lawrence Summers, 2015, Inflation and Activity - Two Explorations and their Monetary Policy Implications, *NBER Working Paper*.

Blanchard, Olivier J., 1993, Suggestions for a New Set of Fiscal lndicators, *The Political Economy of Government Debt*. H. A. A. Verbon and F. A. A. M. Van Winden. (Amsterdam: Elsevier Science).

Blanchard, Olivier J., and Daniel Leigh, 2013, Growth forecast errors and fiscal multipliers, National Bureau of Economic Research.

Blanchard, Olivier J., and Lawrence H. Summers, 1986, Hysteresis and the European unemployment problem, *NBER Macroeconomics Annual 1986, Volume 1* (MIT Press).

Carnot, Nicolas, and Francisco de Castro Fernández, 2015, The discretionary fiscal effort: an assessment of fiscal policy and its output effect. European Economy, European Commission.

Claeys, Grégory, Zsolt M. Darvas, and Alvaro Leandro, 2016, A proposal to revive the European fiscal framework, Bruegel Policy Contribution.

Coibion, Olivier, Yuriy Gorodnichenko, and Mauricio Ulate, 2017, The Cyclical Sensitivity in Estimates of Potential Output. Working Paper, National Bureau of Economic Research.

Coibion, Olivier, Yuriy Gorodnichenko, and Mauricio Ulate, 2018, Real-Time Estimates of Potential GDP: Should the Fed Really be Bitting the Brakes? Policy Futures, Center on Budget and Policy Priorities.

DeLong, J. Bradford, and Lawrence H. Summers, 2012, Fiscal Policy in a Depressed Economy [with Comments and Discussion], *Brookings Papers on Economic Activity*, 233–297.

European Commission, 2005, Public Finances in EMU 2005, *European Economy*.

European Commission, 2014, Assessing public debt sustainability in EU member states: a guide. Occasional Papers, European Commission.

Fatás, Antonio, 2000, Do business cycles cast long shadows? Short-run persistence and economic growth, *Journal of Economic Growth* 5, 147–162.

Fatas, Antonio, and Ilian Mihov, 2009, Fiscal Policy, *EMU at Ten: Should Sweden, Denmark and the UK Join* (SNS Economic Policy Group).

Fatás, Antonio, and Ilian Mihov, 2013, Recoveries, *CEPR Discussion Papers*.

Fatas, Antonio, and Lawrence H. Summers, 2017, The permanent effects of fiscal consolidations, *Journal of International Economics*.

Friedman, M., 1993, The plucking model of business fluctuations revisited, *Economic Inquiry* 31, 171–177.

Gechert, Sebastian, Gustav A. Horn, Christoph Paetz, and others, 2017, Long-term effects of fiscal stimulus and austerity in Europe.

House, Christopher L., Linda L. Tesar, and Christian Pröbsting, 2017, Austerity in the aftermath of the great recession. NBER Working Paper 23417, NBER.

IEO, 2014, An Assessment of IMF Medium Term Forecasts of GDP Growth, (Independent Evaluation Office (IEO) of the International Monetary Fund).

IMF, 2015, Where are we headed? Perspectives on Potential Output, *World Economic Outlook* Chapter 3.

International Monetary Fund, 2011, Modernizing the Framework for Fiscal Policy and Public Debt Sustainability Analysis, .

Irish Fiscal Advisory Council, 2015, The EU Expenditure Benchmark: Operational Issues for Ireland in 2016, *Analytical Note* 7.

Jordà, Òscar, and Alan M. Taylor, 2016, The time for austerity: estimating the average treatment effect of fiscal policy, *The Economic Journal* 126, 219–255.

Kuang, Pei, and Kaushik Mitra, 2018, Potential Output Pessimism and Austerity in the European Union, .

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* 239, R14–R31.

Mc Morrow, Kieran, Werner Roeger, and Valerie Vandermeulen, 2017, Evaluating Medium Term Forecasting Methods and their Implications for EU Output Gap Calculations. Discussion Paper, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Neftçi, Salih N., 1984, Are Economic Time Series Asymmetric over the Business Cycle?, *Journal of Political Economy* 92, 307–328.

Rawdanowicz, \Lukasz, Romain Bouis, Kei-Ichiro Inaba, and Ane Kathrine Christensen, 2014, Secular Stagnation: Evidence and implications for economic policy, *OECD*.

Stadler, George W., 1990, Business cycle models with endogenous technology, *The American Economic Review*, 763–778.

Turrini, Alessandro, and Martin Larch, 2009, The Cyclically-Adjusted Budget Balance in EU Fiscal Policy Making: A Love at First Sight Turned into a Mature Relationship, .

Appendix A. Calculating forecast errors for potential and actual GDP.

When it comes to GDP forecast errors our methodology is straightforward. Let the value of a variable in year t of a variable Y for the year $t + i$ be

$$Y_{t+i}^{t}$$

If $i > 0$ we are talking about a forecast. So for GDP in 2009, the forecast made in 2007 is expressed as

$$GDP_{2009}^{2007}$$

We can compare these forecasts with the actual data for GDP at a later date to compute the forecast error. In some cases when we are talking about a future date from the perspective of both years we are calculating the change in forecast between the two years.

As an example, we can calculate the forecast error for the year 2009 made in 2007 by comparing to the data from the 2018 vintage of the WEO as:

$$FE_{GDP,2009}^{2007-2018} = gdp_{2009}^{2018} - gdp_{2009}^{2007}$$

Where small letters denote logarithms. The only issue we face when comparing these two GDP levels is that because of data revisions, changes in base year and also changes in national accounting rules, the forecast and the actual data might not be comparable as they might not be in the same units or follow the same national accounting criteria.

Because we are interested in revisions to growth rates, we make the two numbers comparable by rebasing the original WEO 2007 real GDP series and its forecasts so that the 2006 data matches the data for that year of the WEO April 2018. In other words, given that the 2006 data now coincides in both the April 2007 and the April 2018 databases, the expression above can simply be calculated as the forecast error of accumulated GDP growth from 2006 to 2009:

$$FE_{GDP,2009}^{2007-2018} = (gdp_{2009}^{2018} - gdp_{2006}^{2018}) - (gdp_{2009}^{2007} - gdp_{2006}^{2007})$$

When it comes to potential output we face a more challenging task. In April 2007 when the IMF or the European Union produce a number for potential output for 2006, this is not observed, it is an estimation of what they believed at that point potential was. Future values of potential output are also dependent on their views at that point in time.

Later when the crisis is in full force we change our views of potential output, but we also change our views on the level of potential output in the past. These revisions are very large, and they completely change the perception of potential output levels in previous years. This means that a calculation of forecast errors of the level of potential GDP based on the accumulation of forecast errors of potential growth rates, as calculated above, would be misleading. Because we have dramatically changed our views on potential output for both the current and future years by rewriting history, it would seem as if the previous path of potential output (measured in growth rates) has not changed that much. But the level has, and we need to incorporate that in our estimates.

To properly calculate the forecast error

$$FE_{POT,2009}^{2007-2018} = (pot_{2009}^{2018} - potr_{2006}^{2018}) - (pot_{2009}^{2007} - pot_{2006}^{2007})$$

Where small letters denote logarithms and $potr_{2006}^{2018}$ is the rebased estimate of potential output for 2006 using the GDP known in 2018 but adjusting it for the output gap as seen in 2007

$$potr_{2006}^{2018} = \ln (GDP_{2006}^{2018}/GAP_{2006}^{2007})$$

Where GAP_{2006}^{2007} is the output gap estimated for 2006 in the 2007 vintage

$$GAP_{2006}^{2007} = \frac{GDP_{2006}^{2007}}{POT_{2006}^{2007}}$$

Appendix B. List of Countries.

AMECO	IMF WEO
Austria	Austria
Belgium	Belgium
Bulgaria	Canada
Croatia	Cyprus
Cyprus	Czech Republic
Czech Republic	Denmark
Denmark	Estonia
Estonia	Finland
Finland	France
France	Germany
Germany	Greece
Greece	Hong Kong
Hungary	Iceland
Ireland	Ireland
Italy	Israel
Latvia	Italy
Lithuania	Japan
Luxembourg	Korea
Malta	Latvia
Netherlands	Luxembourg
Poland	Malta
Portugal	Netherlands
Romania	New Zealand
Slovakia	Norway
Slovenia	Portugal
Spain	Singapore
Sweden	Slovak Republic
United Kingdom	Slovenia
	Spain
	Sweden
	Switzerland
	Taiwan
	United Kingdom
	United States