Fiscal Policy, Potential Output and the Shifting Goalposts

Antonio Fatás
INSEAD, CEPR and ABFER

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. Contractionary fiscal policy, via hysteresis effects, caused a reduction in potential output that validated the original pessimistic forecasts and led to a second round of fiscal consolidation. The evidence suggests that this succession of contractionary fiscal policies was likely self-defeating for many European countries as 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. I would like to thank the discussant, the editors, two anonymous referees as well as the audience for their helpful comments.

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?

Similar questions were asked ten years ago when celebrating the first decade of the Euro. In the Fall of 2008 I contributed (Fatas and Mihov (2009)) to a review of the first 10 years of the Euro, where the overall conclusion was positive: The Euro had been a success and defeated the pessimistic scenarios that some had predicted.

Despite this optimistic view, there was also a warning about unfinished business when it came to fiscal policy. In particular, there was great uncertainty about what could happened if the Euro area had to face a significant downturn. So far, the economic conditions of the first years of the Euro had been benign, years of relative calm from a macroeconomic point of view except for a small and synchronized crisis in the years 2002-2003. Was fiscal policy equipped to address a more volatile environment? The answer in Fatas and Mihov (2009) was no. 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, continued to be ill-designed to handle a large asymmetric shock.

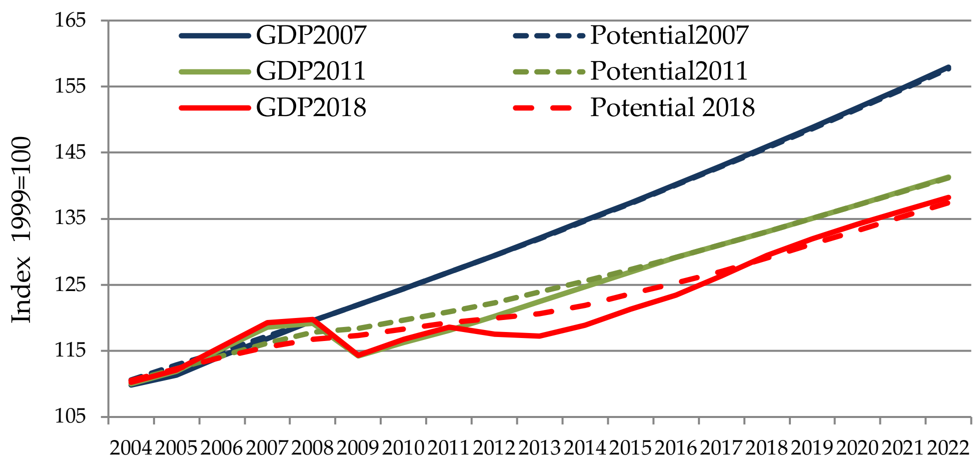
The warning was clear: *"one cannot ignore that these results also portray a picture of a failing institutional framework for fiscal policy. One that has been focused on the concepts of sustainability and coordination without much success, at a time where it should have been focused on* ***strengthening automatic stabilizers and designing a proper framework for the conduct of fiscal policy over the business cycle****."* (Fatas and Mihov (2009))

Unfortunately, this warning and its associated risks became evident as the article 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 and fiscal policy, in particular after

2010, becoming highly procyclical as a reaction to the crisis and the associated 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.[1] GDP forecasts were downgraded several times during this period.

Figure 1. Revisions to Euro actual and potential GDP



The revisions were significant in the first years, but they continued even after 2012 as the Euro zone entered its second recession. And when GDP forecasts were being revised downwards, they also did so for long horizons. 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 started because of exploding deficits and debt. The conclusion was that a process of fiscal consolidation was

[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.

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 contractionary fiscal policy played a big role in those years and GDP growth suffered as a result (Blanchard and Leigh (2013)).

But the damaging effects of fiscal policy did not stop there. The effects of fiscal policy on output fed into more pessimistic views on the future and triggered additional fiscal consolidations. These effects spread via estimates of potential output that were highly procyclical. When designing fiscal policy, in particular when sustainability is an issue, governments need to have an accurate assessment of potential GDP and the output gap.[2] If cyclical events lead to immediate reductions to long-term projections of GDP, it might lead to even more contractionary fiscal policy and further negative effects on output.

These negative effects on output can become permanent via hysteresis effects. Fatas and Summers (2018) provides evidence of such effects during the fiscal contraction of 2010-2011 in Europe. Countries that implemented stronger consolidations saw a much larger reduction in long-term GDP as well as potential output.[3] In this paper we take these results even further and argue that, 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 that we call the "fiscal policy doom loop". Low GDP growth is erroneously 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. As the effects get magnified by this vicious cycle, they become large enough so that we end up with self-defeating fiscal contractions: while policy makers' goal is to improve sustainability, debt-to-GDP ratios end up higher than before, because of the negative effects on 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-

[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).

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 provide the necessary stabilization policy in the presence of cyclical fluctuations. To ensure both goals, policy makers need a long-term forecast of GDP, which normally is expressed as an estimate of current and future potential GDP. In the case of the sustainability assessment, long-term GDP growth rates feed into tax revenues and they also mechanically affect 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 necessary to design countercyclical budgets.

The medium-term objective

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.

Sustainability can be characterized as a steady-state level of the primary balance that ensures a stable path for government debt. 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 an estimate 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 government spending changing proportionally to potential GDP.

$$\Delta G_t = \tau \, \Delta \, Y_t^p$$

In practice this means that government spending is growing (percentage wise) at a rate very close to the growth rate of potential GDP, given that τ is similar to government size ($\frac{G_t}{Y_t}$).[4]

Stabilization policy when output deviates from potential

Under the assumption that fiscal policy is a useful stabilization tool, budget balances must react to cyclical conditions, where cyclical conditions are captured by deviations of GDP from potential output. What is the right cyclical fiscal policy stance and how do we measure it? Following Blanchard (1993) we can use the budget balance as a ratio to GDP as a proxy for the aggregate demand effect of fiscal policy in a given year. [5]

[4] The closer the interest rate is to the growth rate of GDP of the smaller is the initial level of debt, the closer the two numbers will be.

[5] This is the right measure of fiscal policy stance 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.

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 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:

$$CAB_t = T(Y_t^p) - G(Y_t^p)$$

where $T(\cdot)$ and $G(\cdot)$ represent how revenues and spending depend on the level of economic activity.[6]

We can then 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).[7] 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.

Reacting to permanent shocks

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.[8]

---

[6] 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 $T_t^{CAB} = T(Y_t^p) = T_t\left(\frac{Y_t^p}{Y_t}\right)^{\epsilon_T}$ and $G_t^{CAB} = G(Y_t^p) = G_t\left(\frac{Y_t^p}{Y_t}\right)^{\epsilon_G}$
where ϵ_T and ϵ_G are the elasticities of taxes and spending relative to potential output.

[7] The expression for automatic stabilizers is

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

[8] 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

Imagine the government forecasts a change in potential output for this year ($\Delta\, Y_{F,t}^{p}$). Assuming the fiscal adjustment happens in anticipation of this change and takes place via government spending, we should see government spending changing relative to plan by an amount

$$\Delta G_t = \tau\, \Delta Y_{F,t}^{p}$$

This ensures that the path of debt changes by the same amount as the path of potential output and the ratio stays constant. If the government had not anticipated the shock, then after a permanent shock to GDP it would require an adjustment to structural balances (as the debt-to-GDP ratio would be higher than planned).

Reacting to cyclical shocks

Let's start with the budget balance at the medium-term objective level. In the presence of a cyclical negative shock, a recession, the budget balance declines because of automatic stabilizers. The structural balance will not change unless discretionary policies are being implemented. The effects of automatic stabilizers will disappear over time and budget balances will return naturally to its medium-term target as the recessionary effects die out. If, in addition, there was a countercyclical discretionary change in the budget balance, there will be a need to reverse those expansionary policies assuming we want to return to the same debt-to-GDP level. The size of the adjustment will depend on the size of the stimulus implemented as well as the other parameters (such as the interest rate).[9] 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.

When governments mistake cyclical as permanent shocks: procyclical fiscal policy

When growth fluctuates, policy makers must distinguish permanent and cyclical fluctuations in order to decide how much stabilization to provide and by how much we need to adjust fiscal policy to send it back to a sustainable path.

---

[9] In the case of automatic stabilizers, and in the presence of a negative shock, the cyclical decrease in the budget balance increases debt and might also trigger a need to readjust future budgetary plans, assuming that the previous level of debt is considered optimal.

If cyclical fluctuations are mistaken as permanent ones, fiscal policy will become procyclical. To illustrate this point we consider a very simple scenario: an anticipated shock that is temporary in nature but that governments mistake for a permanent one. Authorities will strive to keeping both the budget balance and the structural balance constant by adjusting spending.

Let $\Delta Y_{F,t}^{p}$ be the policy makers' forecast of the permanent shock. Then, governments will adjust government spending by an amount

$$\Delta G_t = \tau\, \Delta\, Y_{F,t}^{p}$$

Because the change in output is cyclical and not permanent, this adjustment means that fiscal policy is procyclical. The budget balance should have gone into a deficit instead of staying constant.[10]

What are the consequences of procyclical fiscal policy? It has an effect on GDP growth via the standard multiplier, and we expect output to change by an amount

$$\Delta Y_t = \mu\, \Delta G_t$$

Where μ represents the short-run multiplier. This means that fiscal policy is not stabilizing output, on the contrary, it is making output more volatile.

Procyclicality can become persistent over time and generate additional rounds of fiscal adjustment (the "fiscal policy negative loop")

After the fiscal consolidation has been implemented, several forces can lead to persistent procyclicality via additional fiscal policy errors.

First, if policy makers ignore the effects of their policies or they underestimate fiscal policy multipliers (μ), they will be surprised by the decrease in GDP. They were expecting a change in GDP equal to $\Delta\, Y_{F,t}^{p}$ but instead they observe $(1+\mu)\Delta Y_{F,t}^{p}$ because of the additional effects of the fiscal consolidation. This decrease in growth rates worsens budget balances and it can lead policy makers to believe again that some of this decrease is permanent, which will require an

---

[10] This procyclicality means an increase in the structural balance, if measured using the correct value for potential GDP. Governments are unaware of this procyclical stance because they are using the wrong value for potential.

additional fiscal adjustment to restore sustainability. Fiscal policy would now be even more procyclical and further away from its optimum.

In addition, if policy makers treat the structural balance not just a target, but also as an indicator of how sustainable fiscal policy is, mismeasurement of the tax elasticities could create additional procyclicality in fiscal policy. This is likely to happen when the elasticities used to calculate cyclically-adjusted variables vary themselves with the cycle. A mismeasurement in the tax or spending elasticities can show a declining structural balance, even if only automatic stabilizers are operating. This will add another reason for an additional contraction in fiscal.

The fact that fiscal policy actions caused by erroneous beliefs on the policy makers' assessment of potential output can lead to further fiscal policy errors creates a "fiscal policy negative loop". But notice that, so far, while these effects can become persistent, they are transitory in nature. Over time, GDP will return to its trend and the policy makers' pessimism about potential output will be slowly corrected. The cost of fiscal policy errors will simply be increasing GDP volatility but no change in GDP growth rates or long-term GDP levels.[11]

Potential output and fiscal consolidation in the presence of hysteresis: the "fiscal policy doom loop"

We now open the door for much larger negative effects of procyclical policy 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 temporarily affected by cyclical conditions, then hysteresis is present across a variety of macroeconomic models (Stadler (1990), Fatás (2000)). In this context hysteresis should simply be interpreted as the years of growth that are permanently lost. The literature has recently regained interest because of the persistence of GDP after the 2008 crisis (Rawdanowicz et al. (2014)).

In addition, the presence of hysteresis makes the assessment and forecasting of potential output much harder. Potential output is affected by cyclical conditions, including those related to fiscal policy decisions. This means that an error in

[11] As long as shocks to GDP are symmetric and the mismeasurement of potential output or the output gap are also symmetric.

fiscal policy now leads to permanent effects. And these permanent effects will potentially validate the erroneous pessimistic forecasts about potential output.

To illustrate these dynamics, we start with the same scenario where a crisis is forecasted for next year. The crisis is cyclical in nature, but the government misinterprets it as a change in potential output. As a result, it decides to engage in a fiscal consolidation via a one-year decrease in spending relative to its normal path (ΔG_t)

The change in spending affect negatively GDP through the fiscal policy multiplier (μ), as discussed earlier.

$$\Delta Y_t^c = \mu \, \Delta G_t$$

Where we use the superscript c to refer to the cyclical effects of fiscal policy.

We now introduce hysteresis by postulating 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) by a factor η that we call the hysteresis parameter.

$$\Delta Y_t^p = \eta \, \Delta Y_t^c = \eta \, \mu \, \Delta G_t$$

The obvious consequence of hysteresis is that policy mistakes now have permanent effects on output. Procyclical policy is now much more damaging. The same logic applies to any other mistake in assessing the state of the economy, these errors will now be magnified and their costs will become permanent. For example, take the case of a government that underestimates the short-run multiplier (μ). In response to a fiscal policy contraction, not only output will be unexpectedly lower (as before) but now potential output will be affected as well in a way that will surprise policy makers. This means that the initial pessimism on long-term GDP might even worsen over time, leading to additional rounds of fiscal consolidation. We call this the "fiscal policy doom loop".

This loop between unfounded pessimism, contractionary fiscal policy and negative effects on potential output can be difficult to break because the permanent consequences of fiscal policy serve as a false confirmation of the initially unfounded pessimistic forecasts of policy makers. Policy makers might feel that their original pessimism was correct when GDP ends up in a lower trajectory than before the crisis. They fail to realize that part of what they see is the effects of the policies that were implemented because of their initially-wrong expectations about long-term growth.

Finally, there is also the possibility that the overall effects are stronger when we consider the way in which policy makers learn over time about the nature of the shocks affecting the economy. If governments are initially too pessimistic and confusing a cyclical change with a change in potential, it should be the case that as time passes they learn about their mistake. But given that potential is an estimate of long-term GDP, it might take years for governments to realize of their pessimism. In the meantime, weak temporary growth will provide additional negative signals that might strengthen their pessimism. During these years, their procyclical policies (combined with the other cyclical forces) will lead to a spiral of reductions in growth that will partly be interpreted as permanent and will deteriorate further the debt sustainability analysis of governments. As time passes, hysteresis effects make these cyclical GDP changes turn into permanent, eliminating the possibility of a recovery that would reverse the original mistaken estimates of potential output.

In the presence of hysteresis, procyclical fiscal policy now has stronger negative effects on welfare. In addition, it can even be that its economic outcome is the opposite of what was intended. Fiscal consolidations are implemented to reduce government borrowing in order to bring the debt-to-GDP ratio under control. But given the negative effects on GDP, fiscal consolidations will have a negative effect on the debt-to-GDP ratios through hysteresis. If short-run fiscal multipliers and hysteresis effects are large, it is very likely that fiscal consolidations will be self-defeating. DeLong and Summers (2012) provide supporting evidence for this possibility in the US economy and Fatas and Summers (2018) provide similar evidence for the European economies during the 2010-11 fiscal consolidation episode.

An alternative scenario: hysteresis after a fiscal consolidation

In our previous example, we had chosen a scenario where we started with policy makers mistakenly assuming that a cyclical event was permanent. There is no reason for this to be the starting point. Imagine an economic crisis that leads to higher debt because of a succession of deficits triggered by automatic stabilizers or because of support provided to the financial system. Governments engage in a fiscal consolidation either because of the need to bring debt back to normal or because they are forced by the market that pushes interest rates on government debt to higher levels.

Fiscal consolidation will lead to an economic downturn that through hysteresis will affect potential output. These downward revisions of potential will be even larger if governments underestimate fiscal policy multipliers or the hysteresis parameter. These revisions will require additional tightening of fiscal policy. In this scenario we also observe the same "doom loop" between fiscal policy and revisions to potential output estimates even if there was no mistake in the *initial* expectations of policy makers.

It is very likely that in the period 2008-14 we saw a combination of both of these dynamics: initial pessimism plus the need to implement consolidations because of a large crisis. These dynamics reinforce each other and produce large fiscal consolidations with significant effects on output. We will explore this behavior and effects in our empirical section.

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. Over the years, the framework 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.[12] As a result, it introduced the production of a regular Debt Sustainability Monitor report (twice a year) with a detailed government debt sustainability exercise.[13] These reports followed the standard logic of debt

[12] See European Commission (2005).

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

sustainability and they are broadly consistent with the methods used by other organizations such as the IMF.[14]

It was the same 2005 reform that made explicitly the role of economic factors in justifying deviations from the 3% limit.[15] 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.

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 ten-year window with estimates of the past five years combined with European Commission forecasts for the next four. 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 two years are combined with the trends estimated by the production function. The trend produces a five-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).

---

[14] The revised framework of the IMF was introduced in International Monetary Fund (2011).
[15] 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)

Several assessments of these methodologies have reached the conclusion that potential output estimates react too strongly to cyclical changes in GDP. They tend to be too optimistic during good years and too pessimistic during bad years (Mc Morrow, Roeger, and Vandermeulen (2017) or Kuusi (2017)). In all these assessments the cost of procyclical potential output estimates is seen as procyclical fiscal policy and more volatile GDP. As we have argued in the previous section, and in the presence of hysteresis, this procyclical policies can create more damaging long-term effects on GDP. And these hysteresis effects might wrongly validate the unfounded forecasts on which policy was based. 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, potential GDP in the Euro area remains about 15% lower than its pre-crisis trend.

In this section we explore how this revision to long-term forecasts happened and 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 three sources.[16] Our main source of data is vintages of the IMF World Economic Outlook (WEO) during the years 2007 to 2018.[17] For each year we focus on the April issue of the WEO and collect data for the group of advanced economies, although our focus will be on European countries, in particular Euro members. Each of these vintages provides 5-year forecasts for GDP as well as potential GDP. We take the

---

[16] Countries included in the three sources are listed in a data appendix.

[17] 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.

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 check some of our results using 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 choose 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.

Finally, we also check whether the general pattern we discover is consistent with similar (one-year) forecasts by the OECD.[18] We restrict our sample to countries that are in the group of advanced economies, as defined by the IMF, and select the Spring forecasts to match the other two sources.

Measuring forecasts errors

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 any of our three sources and look at a two-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:

[18] Data was kindly provided by Thomas Chalaux from their work Chalaux and Guillemette (2018)

$$FE^{2010,2012}_{GDP,2011} = gdp^{2012}_{2011} - gdp^{2010}_{2011}$$

Where small letters represent logarithms. Notice that GDP^{2012}_{2011} 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.[19] One way to avoid data revisions of GDP levels is to rewrite the expression for forecast errors in terms of GDP *growth*.[20] 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

$$FE^{2007}_{GDP,2009} = (gdp^{2012}_{2011} - gdp^{2012}_{2009}) - (gdp^{2010}_{2011} - gdp^{2010}_{2009})$$

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 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 based on 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.[21] 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

---

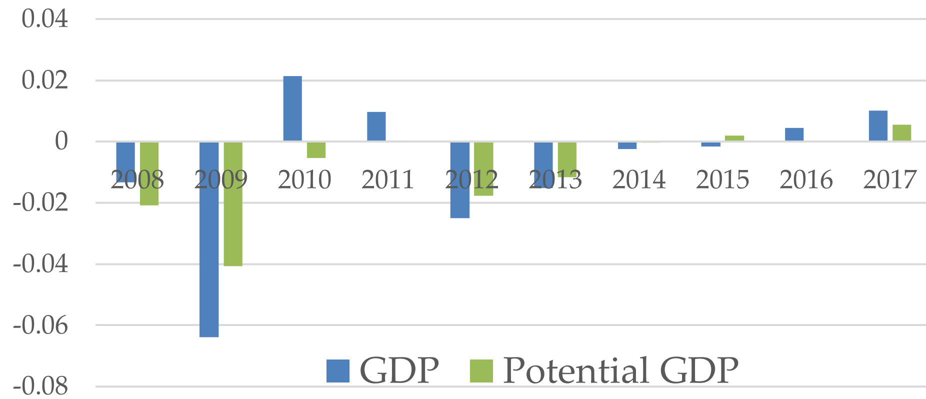
[19] For example, since October 2014 the WEO has started using updated data using ESA2010 criteria.

[20] 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.

[21] 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.

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 window 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 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.

The procyclical nature of potential output estimates

To highlight this correlation between revisions to potential GDP and GDP surprises we calculate one-year, two-year and six-year forecast errors for a sample of all advanced economies for which data is available in the WEO. We then do the same one-year and two-year calculations for the sample of European countries in AMECO as well as the OECD database.

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 two-year and six-year forecasts, we only pick non-overlapping years.[22] In all our regressions we include country fixed effects as well as year dummies.[23]

This 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.

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

1-Year Forecast Error Potential GDP						
	IMF WEO			OECD		
	Advanced	Europe	Euro	Advanced	Europe	Euro
	(1)	(2)	(3)	(4)	(5)	(6)
1-Year Forecast	0.648***	0.671***	0.735***	0.815**	0.934***	1.176**
Error GDP	(0.0446)	(0.0458)	(0.0917)	(0.312)	(0.327)	(0.440)
Constant	-0.0050***	-0.0064***	-0.0049***	-0.0049**	-0.0035**	-0.0063*
	(0.0003)	(0.0004)	(0.0004)	(0.0022)	(0.0016)	(0.0031)
Observations	370	269	185	258	197	134
R-squared	0.555	0.534	0.657	0.213	0.238	0.279
Countries	35	26	18	30	23	16

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

Table 1 shows the result for three separate samples: Advanced, Europe and the Euro area for the IMF and OECD databases. [24] There is strong consistency across the five columns. there is a very strong relationship between the two variables and the coefficient is around 0.65-1.10, signaling that surprises in GDP lead to large immediate revisions to our long-term views on output. We replicate the same analysis but for the AMECO database in Table 2 and we obtain similar results.

[22] 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.
[23] Results are practically identical if we remove year dummies or fixed effects. Clustered standard errors in parentheses using Driscoll and Kraay (1998).
[24] 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.

Table 2. Revisions to Potential GDP (AMECO)

	1-Year Forecast Error Potential GDP	
	AMECO	
	Europe (1)	Euro (2)
1-Year Forecast Error GDP	0.853*** (0.0733)	0.870*** (0.0728)
Constant	0.0064*** (0.0097)	0.0064*** (0.0063)
Observations	232	152
R-squared	0.704	0.722
Countries	30	19

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

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 3 replicate present the results.[25]

There is a small 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 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 do that for the IMF

[25] We focus on the Europe and Euro samples (results for the group of advanced economies are very similar). We do not include the OECD database in the table for the sake of space. Results are very similar to those of the AMECO database, there is no significant asymmetry in the way potential GDP is revised in response to GDP growth surprises.

and AMECO databases.[26] In the case of the AMECO database the longest window is two years. In the case of the IMF database we can have a six-year window. 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 4.

Table 3. 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	0.496***	0.553***	0.173	0.837***	0.851***
Error GDP+	(0.102)	(0.128)	(0.185)	(0.0724)	(0.0702)
1-Year Forecast	0.890***	0.914***	1.015***	0.961***	1.021***
Error GDP-	(0.125)	(0.153)	(0.146)	(0.0923)	(0.120)
Constant	-0.0050***	-0.0039***	-0.0014	0.0062***	0.0036**
	(0.0010)	(0.0008)	(0.0010)	(0.0009)	(0.0063)
Observations	269	185	113	232	152
R-squared	0.544	0.664	0.625	0.706	0.725
Countries	26	18	18	30	19

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

Overall, we confirm our previous results, except that 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 coefficients on the two-year or the six-year windows are not too different from the one-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.

[26] The AMECO database is interesting because it reflects the views of the European economic authorities when judging the appropriateness of fiscal policy plans. The IMF database is the only one that includes 5 year forecasts so it allows us to analyze a longer horizon.

Table 4. 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	0.823***	0.877***	0.945***	0.812***	0.856***
GDP	(0.0253)	(0.0174)	(0.000)	(0.0160)	(0.013)
Constant	0.0080***	0.0032***	0.0163***	-0.0086***	-0.0001
	(0.0016)	(0.0004)	(0.0035)	(0.0001)	(0.00009)
Observations	123	85	31	86	57
R-squared	0.868	0.912	0.995	0.706	0.760
Countries	26	18	17	30	19

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

We provide an additional robustness check in Table 5, 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 5. Alternative measures of Potential GDP

2-Year Forecast Error Potential GDP		
	AMECO (Euro)	
	Trend GDP	Uncorrected Potential
	(1)	(2)
Forecast Error	0.799***	0.356***
GDP	(0.000)	(0.0913)
Constant	0.0054***	-0.00079
	(0.00310)	(0.00063)
Observations	38	57
R-squared	0.649	0.330
Countries	19	19

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

In column (2) we also 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. As we argued earlier, 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.[27]

5. EU Fiscal Policy During 2007-2014: The Role of 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.

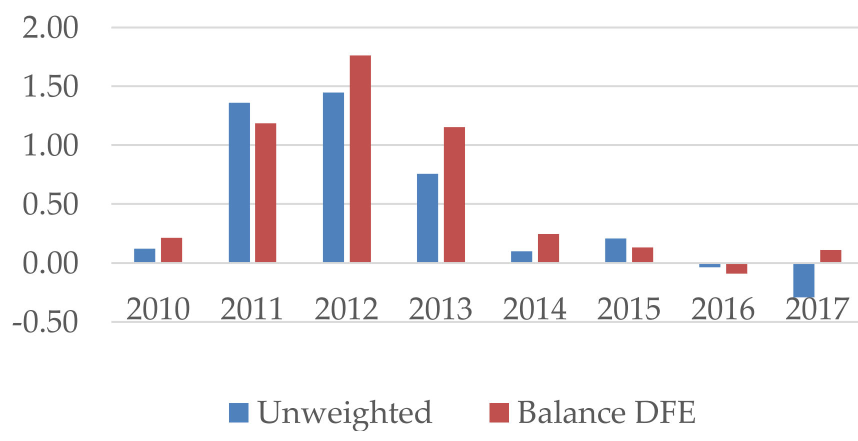
In Figure 3 we show how fiscal plans evolved during these years. We show one-year forecast of changes in structural balances (measured as % of GDP) for the average of Euro members. We calculate the change both as an unweighted average and one where countries are weighted by their GDP size in 2008. The 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.

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

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) indicator.[28]

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

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



[28] 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). 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.

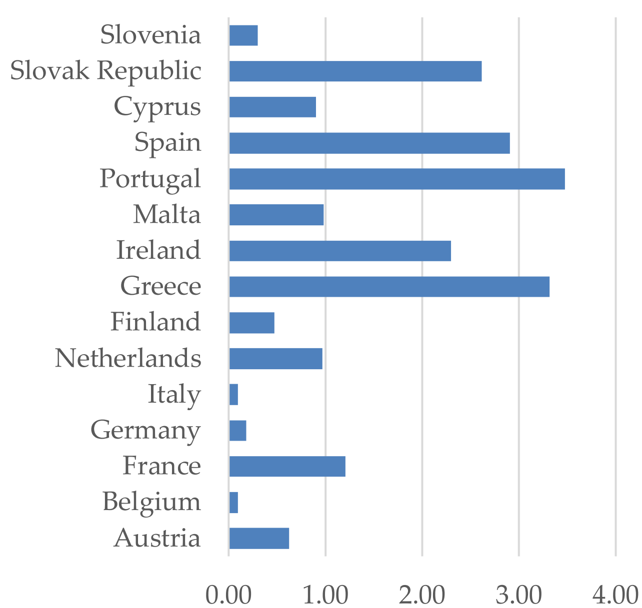
Figure 4 compares the two magnitudes and it is clear that the overall numbers are very similar even if there are differences because of their methodologies (Carnot and de Castro Fernández (2015)). We see a wave of fiscal consolidation that starts in 2010 and continues until 2013. We will now study these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

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 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.

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.

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



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 ($\Delta SB_{i,2010-2011}^{2010}$) as an indicator of the fiscal consolidation. We then regress the forecast error for real GDP for those two years on this measure.[29]

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

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 *real-time* 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

[29] 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 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.

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, consistent with our 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.

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 (2018) confirms that fiscal consolidation had an immediate effect on our long-term views on output.[30]

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

30 In Fatas and Summers (2018) as well as Gechert et al. (2017) there is a variety of robustness tests over different horizons that confirm this result.

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 surprise change in inflation will have an effect on the real value of the debt (which is denominated in nominal terms).

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.[31]

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

[31] 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.

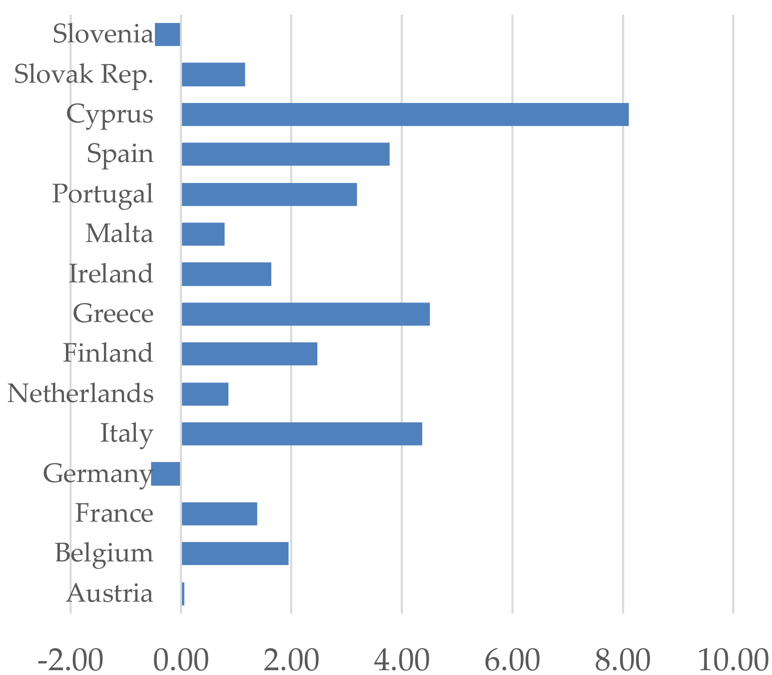
appear in the April 2012 vintage. How did our plans for those years changed between April 2010 and April 2012?

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.

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)).

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



We now present a 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 potentially potential effects, pushed governments to a much tighter and faster response to changing economic 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)	(0.0298)
Constant	0.0143***	0.0225***	0.00275***
	(0.00505)	(0.00582)	(0.000739)
Observations	23	15	146
R-squared	0.059	0.099	0.364
Number of countries	23	15	18

Robust standard errors in parentheses (Clustered for Column (3))
*** p<0.01, ** p<0.05, * p<0.1

In Table 10 we run the same regression but using surprises in potential output as the explanatory variable. We run a 2SLS specification where we are instrumenting for changes in potential output using the changes to GDP during the same years. In other words, we are testing how the changes in long-term forecasts that originated in the surprises to GDP during the 2009-2011 period modified our fiscal plans. We do that using changes in both real and nominal potential GDP growth given that nominal changes matter for debt sustainability.

The coefficients are significant and larger for the case of nominal changes in potential GDP, as we would expect given our previous results.

Table 10. Changing Economic Conditions and Future Fiscal Plans (2SLS)

	Europe 2011-2013	Euro 2011-2013	Europe 2011-2013	Euro 2011-2013
	(1)	(2)	(3)	(4)
Forecast Error Potential GDP (previous 2 years)	-0.343*	-0.334*		
	(0.167)	(0.173)		
Forecast Error Pot. GDP Nominal (previous 2 years)			-0.549**	-0.582**
			(0.241)	(0.249)
Constant	0.0119**	0.0207***	0.0119**	0.0190***
	(0.00505)	(0.00627)	(0.00475)	(0.00619)
R-squared	--	--	0.052	0.020
Number of countries	23	15	23	15

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

The size of this adjustment is not far from what our static example of Section 2 suggested. In that case, to keep debt stable, the budget balance had to be adjusted by a factor of τ relative to the surprise in potential output. In our example, τ was both the average and the marginal tax rate. Average tax-to-GDP ratios for our sample will be around 0.40 and marginal tax rates will be higher. These numbers are not far from the coefficients above.

Table 9 and Table 10 have confirmed that worsening economic conditions led to a tightening of fiscal policy plans for the years 2012-13. But were these plans implemented? What type of policies did we see during those years? We make use of the Discretionary Fiscal Effort (DFE) indicator from AMECO to check the actual tightening of fiscal policy in those two years. These are actual implemented changes in tax rates and spending (expressed as % of GDP). Because this indicator is not available in real time (as a plan) we can simply measure the ex-post assessment of the changes being implemented. We will separate changes on the expenditure and the revenues side.

In Table 11, Using as dependent variable is the DFE (expenditures and revenues) of the years 2012 and 2013 we first run a regression on the GDP surprise for the previous two years (columns (1) and (2)). We see that larger GDP surprises led to more tightening of fiscal policy in the cross-section of countries both on the expenditures and the revenue side. In the next two columns we use potential output as the explanatory variable (with the surprise in GDP as instrumental variable, as we did before). Results are similar. If we add the two coefficients for expenditures and revenues we see an overall effect on the balance that is larger than the one from

Table 9 and Table 10. The total effect on the balance is now around 0.9, larger than the 0.2-0.6 we found earlier. But this is likely to be because some of these policies were already anticipated in the previous years so the actual tightening relative to plans (what we measured in previous tables) was smaller.

Table 11. Actual Fiscal Consolidation 2011-2013 (DFE, AMECO)

	Expenditures	Revenues	Expenditures	Revenues
	OLS 2011-2013	OLS 2011-2013	2SLS 2011-2013	2SLS 2011-2013
	(1)	(2)	(3)	(4)
Forecast Error GDP	0.5285***	-0.3476***		
(previous 2 years)	(0.1374)	(0.0572)		
Forecast Error Pot. GDP			0.6883***	-0.3117***
(previous 2 years)			(0.1572)	(0.0911)
Constant	-0.0223***	0.0196***	-0.01704***	0.01694***
	(0.00344)	(0.00311)	(0.0031)	(0.0036)
R-squared	0.663	0.475	0.634	0.224
Number of countries	17	17	16	16

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

In summary, in the previous three tables we have shown how policy tightened across countries in a way that was related to both the cyclical effects of the previous two years and the downward revisions of potential output. We now explore the effects of this second round of fiscal consolidation?

The cyclical effects of the 2012-13 fiscal consolidation

We first make use of the methodology of Blanchard and Leigh (2013) in Table 12, 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.

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. The first two columns of Table 13 show that the planned fiscal consolidation had a negative effect on our revisions to potential output forecasts that is almost the same size as in the previous consolidation, 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 12) to identify how changes in GDP translated into changes in potential output estimates. The coefficient is close to one also suggesting that this translation was as strong or stronger than during 2010-11.

Table 12. 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

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

Table 13. 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

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

From procyclical forecast errors to long-term GDP effects: the role of hysteresis

Our analysis of the period 2010-14 has revealed a pattern where fiscal policy contractions in 2010-11 led to changes in GDP and downward revisions of estimates of potential output. This pessimism caused additional fiscal consolidations in the years that follow. What was the long-term outcome of these dynamics? Was the initial pessimism justified or did we see policy makers reverse their views on potential output?

We can check the evolution of these forecasts by comparing the 2-year forecast error with the 6-year forecast error from the IMF database using the first fiscal consolidation as an example.

Table 14 shows the result using the April 2010 vintage and regressing the size of the forecast error all the way to 2015 on the size of the forecast error for 2011. 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) were not reversed but it also hints to the fact that something else happened during the following four years that *amplified* the initial pessimism.

Table 14. 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

Our interpretation of these results is that the two rounds of fiscal policy contraction had permanent effects on GDP via hysteresis effects.[32] From a theory point of view, this interpretation is consistent with models where, because of hysteresis, potential output is itself endogenous to cyclical conditions. A recession makes the long-term engines of growth slow down and GDP never returns to its pre-recession trend (Fatás (2000)). There is strong evidence that cyclical dynamics affect potential output through hysteresis effects.[33]

From an empirical point of view, our interpretation is supported by the extensive analysis of Fatas and Summers (2018), who show that the GDP effects of the 2010-11 fiscal consolidation persisted all the way to the 2017 GDP forecasts for the year 2022, a window of more than ten years. Gechert et al. (2017) provide additional supporting evidence that confirms the presence of hysteresis in response to fiscal policy during these years.

Are there alternative explanations? Could it be that the initial pessimism was correct and the permanent effects that we see at the end confirm those pessimistic forecasts? Our results provide evidence that this explanation is not plausible. In 2010 the plans for fiscal consolidation were based on a certain forecast of potential output. We have shown that this forecast was wrong, and the following forecast errors was correlated to the size of the planned fiscal consolidation. In other words, the long-term effects we observe at the end do not simply correspond to the initial pessimism but they are a consequence of the hysteresis effects of the first wave of fiscal consolidation. And it is *these forecast errors* and not the initial pessimism the ones that lead to additional tightening of fiscal policy in the next two years (Table 9 and Table 10).

This interaction between long-term GDP forecasts (potential output), fiscal policy decisions and hysteresis is what we call the "fiscal policy doom loop". Pessimism about the future leads to fiscal policy consolidation. These policies have an effect on GDP that surprises policy makers and causes them to revise even further their views on long-term fiscal policy. As a result, they tighten even further fiscal policy. The cyclical effects of fiscal policy become permanent through hysteresis effects.

---

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

[33] See Blanchard, Cerutti, and Summers (2015), Martin and Wilson (2013), Haltmaier (2013), Rawdanowicz et al. (2014) or Reifschneider, Wascher, and Wilcox (2015) for additional evidence on the long-term effects of cyclical events.

Our dynamics also make it very difficult for policy makers to learn from their mistakes because the permanent effects associated to hysteresis can validate the originally mistaken pessimism. We can do a quick calibration to understand the magnitude of these dynamics using of the framework we developed in Section 2.

Using the example where a cyclical shock is mistaken as a permanent one. $\Delta\, Y^{p}_{F,t}$stand for the anticipation of a permanent shock to GDP. Assuming for simplicity that governments believe that short-run multipliers are zero, governments change government spending by an amount (τ is the average tax rate):

$$\Delta G_t = \tau\, \Delta\, Y^{p}_{F,t}$$

Fiscal policy will affect GDP and some of this will become permanent via hysteresis. The permanent effects of fiscal policy can be expressed as

$$\Delta Y^{p}_{t} = \eta\, \mu\, \Delta G_t = \eta\, \mu\, \tau\, \Delta\, Y^{p}_{F,t}$$

Where μ is the short-run multiplier and η is the hysteresis parameter. The change in potential output coming from fiscal policy depends on three key parameters: hysteresis, the fiscal policy multiplier and government size. We can do a quick calibration of the size of these effects using estimates from previous papers. If we assume fiscal policy multipliers to be around 1.8 (Blanchard and Leigh (2013) or our results above), a value for τ or 40% (an average of tax as a % of GDP for EU countries) and a range for the hysteresis parameter of 0.5-0.7 (see Rawdanowicz et al. (2014) or Fatas and Summers (2018)), the implemented fiscal policies could change potential output by one third to one half of the size of the erroneous forecasts for potential output.[34] In other words, the permanent effects of fiscal policy can already validate half of the mistaken pessimism of policy makers.

But this estimate is a lower bound of the total effects we are likely to see on potential output. Potential output will also change because of hysteresis associated to the cyclical event that caused the crisis. This will again validate pessimistic expectations or, even worse, surprise even more governments and lead to further pessimism and additional rounds of fiscal consolidation. And, as before, policy makers that in addition underestimate fiscal policy multipliers or misread the information contained on structural balances could end up implementing even tighter fiscal policies, further damaging output.

[34] As discussed earlier, the fiscal consolidation in reaction to a surprise in GDP that we observed in 2012-13 is not far to the 40% value we are calibrating in this exercise.

6. Conclusions and policy recommendations.

The doom 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.

Any fiscal framework requires estimates of potential output and it is well known that measurement errors will lead to the wrong fiscal policy.[35] Where there is less clarity is on the consequences and importance of the mismeasurement.[36]

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 a combination of a procyclical bias in the forecasts and possibly an underestimation of fiscal policy multipliers led to a deeper recession.

Our paper goes beyond the idea that procyclical policy generates additional volatility and makes fluctuations more persistent. By considering the role of hysteresis we show that fiscal policy errors get magnified and end up leaving permanent scars on the economy. In addition, those permanent scars have the pernicious effect of generating additional rounds of fiscal consolidation that deepen even further the recession.

One could argue that this hysteresis effect is symmetric and during booms we could see the exact opposite effects. In this case, the actual consequences on the economy would be to add (very persistent) volatility without affecting its long-

---

[35] When we talk about mismeasurement of potential output, we are focusing on the tendency of short-term growth fluctuations to influence long-term GDP forecasts and how it leads to a procyclical bias in fiscal policy. Our paper has nothing to say about the possibility that *on average* forecasts are too optimistic or too pessimistic.

[36] 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.

term path.[37] But this assumes a level of symmetry that is not supported by the data. First, business cycles are not asymmetric. One extreme view is 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 either of these two cases the downside of procyclical policy would be much larger than the upside in good times.

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 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 only 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. In fact, our results are large enough that suggest that fiscal policy decisions were likely to be self-defeating.[38] 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.

The event analyzed in this paper is a unique event: a 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 of 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

[37] 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.

[38] See Fatas and Summers (2018) for a mapping 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.

around it. But this is not correct, 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: Reducing procyclicality

The problem highlighted in this paper starts with the fact that measures of potential output are excessively volatile and highly procyclical. At first glance, this seems like a technical issue that should be improved through better modeling and econometric techniques. However, its interactions with the way policies are set and its consequences considered by the economic analysis remain complex and might require revisions to several aspects of the fiscal policy framework.

The first obvious reaction to our results 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.[39] 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 putting too much weight on recent data to avoid an overaction to cyclical events (Kuang and Mitra (2018)).

There is a second operational aspect of the way in which potential output enters into the design of fiscal policy that could be improved and that has been addressed by recent reform proposals. It has to do with the way we use structural budget balances as indicators of sustainability and as medium run

[39] 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.

targets. 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 more volatile 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".[40] This logic is also supported by the proposals of Bénassy-Quéré et al. (2018) or Claeys, Darvas, and Leandro (2016).

Towards a better fiscal policy framework: Modeling Hysteresis

Smoother estimates of potential output or the use of expenditure rules do not fully address all the concerns raised by our paper. What we also need is a better understanding, in real time, of the nature and persistence of different shocks.

We can potentially rely on economic models to identify shocks and produce better real-time estimates of potential GDP as in Coibion, Gorodnichenko, and Ulate (2018). While this methodology is promising, the results of this paper raise many questions about the economic models that can capture the type of shocks that generate economic fluctuations. The area where traditional models fail is in the characterization of hysteresis, a central feature of our results but absent in standard macroeconomic models.[41]

Eradicating the possibility of the "fiscal policy doom loop" requires designing fiscal policy using economic models that allow for the presence of hysteresis. Recognizing the existence of hysteresis should make economic policies (fiscal

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

[41] Coibion, Gorodnichenko, and Ulate (2018) rely on models that identify supply and demand shocks by assuming that only supply shocks have permanent effects on GDP. This assumption runs contrary to the idea that all types of shocks generate hysteresis and permanent effects.

and monetary) much more aggressive in particular during large negative cyclical events like the one the Euro area experienced during the 2008-2014 period.

Recognizing hysteresis is particularly relevant in the presence of negative shocks given our argument about the inherent asymmetry when it comes to the consequences of procyclical fiscal policy. Having a rule that takes into account this asymmetry, in particular for large shocks, might be challenging and will certainly require a certain amount of flexibility and judgement that does not go well with simple and numerical rules. One way to resolve this trade off is to recognize the need for flexibility during "exceptional circumstances". Such a clause is already present in the EU fiscal policy framework but today this is simply related to a narrow definition of economic circumstances that does not include, for example, the nature of the shock or whether monetary policy can be an effective stabilizing tool its role.

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, Eugenio Cerutti, and Lawrence H. Summers, 2015, Inflation and Activity, . ManuscriptManuscript.

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.

Chalaux, Thomas, and Yvan Guillemette, 2018, If potential output estimates are too cyclical, then OECD estimates have an edge, *OECD ECOSCOPE*.

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.

Driscoll, John C, and Aart C Kraay, 1998, Consistent covariance matrix estimation with spatially dependent panel data, *Review of economics and statistics* 80, 549–560.

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.

Fatás, Antonio, 2000, Endogenous growth and stochastic trends, *Journal of Monetary Economics* 45, 107–128.

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, 2018, The permanent effects of fiscal consolidations, *Journal of International Economics*, 238–250.

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.

Haltmaier, Jane, 2013, Do recessions affect potential output?, *FRB International Finance Discussion Paper* 1066.

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.

Martin, Robert, and Beth Anne Wilson, 2013, Potential Output and Recessions: Are We Fooling Ourselves?, *Manuscript.*

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*.

Reifschneider, Dave, William Wascher, and David Wilcox, 2015, Aggregate supply in the United States: recent developments and implications for the conduct of monetary policy, *IMF Economic Review* 63, 71–109.

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

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

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