# Endogenous Fiscal Consolidations*

Pablo Hernández de Cos  Enrique Moral-Benito

May 2011

**Abstract**

There is evidence in the literature of fiscal consolidation episodes producing (non-Keynesian) expansionary effects (e.g. Alesina and Ardagna, 1998). We replicate this result for a panel of OECD countries under exogeneity of the fiscal tightening decision, and provide evidence that this decision might be endogenous to GDP. Once endogeneity is taken into consideration, we find that expansionary effects disappear. We also investigate the determinants of successful consolidations. In particular, using model averaging to overcome the problem of model uncertainty, we conclude that economic recovery and cuts in public wages are the most appropriate ingredients for successfully reducing budget deficits.

JEL Codes: H30, H62, C23.

Keywords: Fiscal consolidation, panel data, endogeneity, model averaging.

## 1 Introduction

In response to the global crisis that erupted in 2008, comprehensive support packages have been implemented by fiscal authorities in many G20 countries. These expansionary fiscal measures, together with cyclical revenue losses and expenditure hikes, have resulted in sharp increases in budget deficits. Therefore, many governments are already preparing (or have already implemented) budgetary consolidation measures to ensure fiscal sustainability. In this context it is interesting to investigate the potential impact of fiscal consolidations on economic growth, an issue that it is far from having a definitive answer in the literature.

Giavazzi and Pagano (1990) describe the possibility that fiscal consolidation episodes could be expansionary for an economy, challenging the broadly accepted Keynesian notion

*Contact: Bank of Spain. pablo.hernandez_de_cos@bde.es and enrique.moral@gmail.com. We would like to thank Manuel Arellano, Cristian Bartolucci, and seminar participants at the Bank of Spain for useful comments and suggestions. We also thank Silvia Ardagna for kindly sharing the data. The opinions and analyses are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Bank of Spain or the Eurosystem.

concerning the existence of a positive fiscal policy multiplier.[1] In particular, they observed a consumption increase during the fiscal stabilisation in Ireland from 1987 to 1989 and in Denmark from 1983 to 1986. Since this consumption increase was not fully explained by the usual sources such as disposable income, the authors concluded that it was due to the fiscal adjustment and thus these episodes constituted expansionary fiscal adjustments.

Giavazzi and Pagano's (1990) paper has generated a growing literature concerning the so-called non-Keynesian effects of fiscal policy, e.g. Cour et al. (1996), Alesina and Ardagna (1998, 2010), Miller and Russek (2003). These studies are based on empirical analyses in which they first identify periods of drastic and sizeable budget cuts within a panel of OECD countries, and then perform a descriptive analysis of the sample characteristics of macroeconomic aggregates, mainly GDP, before, during, and after the year in which the consolidation episode took place. The main conclusion from this literature is that fiscal adjustments are often followed by an improved growth performance, which is interpreted as evidence of non-Keynesian effects during fiscal consolidation episodes.

A positive correlation between fiscal consolidation episodes and GDP growth does not necessarily mean that fiscal consolidations generate economic growth. In fact, this literature usually assumes that the consolidation episode is exogenous to GDP, and thus causality issues are well beyond the scope of these papers.[2] The positive correlation between fiscal adjustments and economic growth may be the result of a positive effect from GDP growth to fiscal consolidation instead of the other way around as suggested in this literature: the expectation of a recovery (stronger during the trough of the cycle) may increase the likelihood of public finance consolidation (i.e. consolidation episodes are endogenous to GDP). Accepting as valid the empirical specification in this literature, the main aim of this paper is to tackle these endogeneity issues in order to investigate whether there is a causal effect from fiscal consolidation to GDP growth in the short run (i.e. non-Keynesian effects of fiscal adjustment episodes). Our main conclusion is that endogeneity biases might be chiefly responsible for the non-Keynesian results previously found in the literature; hence, fiscal adjustments are found to have a negative effect on GDP growth in the short run.

Another critical point in previous approaches is the definition of the fiscal consolidation episodes. Selecting large-scale fiscal adjustments implies, on the one hand, choosing a measure of fiscal policy, and, on the other, defining what "large-scale" means precisely. As

[1] Feldstein's (1982) paper is probably the first to find evidence in favor of the non-Keynesian hypothesis of fiscal policy. In particular, the paper presents a negative and statistically significant estimate for the public expenditure coefficient in a private consumption function. Feldstein argued that reductions in public expenses may be expansionist if they are seen as an indication of future tax cuts. Kormendi and Meguire (1990) also find evidence of this non-Keynesian result.

[2] Some attempts have been made in the literature to address this potential endogeneity issue. For example, Ardagna (2004) instruments fiscal consolidation episodes with political variables such as the orientation of the government party, but ignores the endogeneity of other fiscal variables such as the size of the consolidation. Giavazzi and Pagano (1996) and Giavazzi et al. (2000) respectively estimate consumption and savings equations accounting for potential endogeneity problems of fiscal variables such as taxes (i.e. the instrument government taxes with the government surplus), but ignore the potential endogeneity of the fiscal adjustment episode per se.

for the measure of fiscal policy, the empirical literature essentially relies on large reductions observed in the cyclically adjusted primary budget balance (e.g. Blanchard, 1993). With respect to the meaning of "large scale", we might consider the size criterion (i.e. a sufficiently large reduction in the primary balance in a given period), the persistence criterion (i.e. a sufficiently long time period during which the primary budget balance constantly improves), or a combination of both. However, all these fiscal consolidation definitions suffer from a potential problem of sample selection. This is so because, according to the different criteria employed in the literature, the different studies only analyse successful consolidations while many failed attempts to reduce fiscal deficits are ignored. In search of exogenous sources of variation in fiscal policy, Ramey and Shapiro (1998), and Romer and Romer (2010) follow a narrative approach for defining large discretionary changes in fiscal policy that do not depend on the success of the policy.[3] Although both papers are based on a VAR framework,[4] this narrative approach seems to be a promising alternative when defining fiscal consolidation episodes. In fact, the IMF's WEO (October 2010) follows this narrative approach and defines fiscal consolidation episodes for a sample of OECD countries over the period 1980-2009. According to the IMF definition, fiscal consolidations are, on average, followed by negative GDP growth in the short run. In this paper we find that, controlling for potential endogeneity biases, the two alternative definitions of fiscal consolidations proposed in the literature provide the same result: fiscal adjustments are found to have a negative effect on GDP growth.

Regardless of their impact on GDP, another crucial issue from a policy perspective is how to succeed in terms of deficit reductions when a fiscal consolidation is carried out. According to the different criteria considered in the literature, a fiscal consolidation is successful if the reduction in the debt-to-GDP ratio is sufficiently large and persistent (e.g. Alesina and Ardagna, 2010; Alesina and Perotti, 1995). The empirical analysis of the main factors driving this success is controversial because there is no theoretical model of reference. The usual approach in the empirical literature investigating the determinants of success in reducing budget deficits is based on a regression of a dummy variable of successful consolidations on a set of candidate determinants (e.g. Alesina and Ardagna, 1998; Giudice et al., 2007). Due to the lack of clear theoretical guidance, the results depend very much on the particular variables included in the regression. In order to empirically overcome this model uncertainty problem, we avoid the model choice problem and consider model averaging techniques. By doing so we estimate all possible models resulting from different combinations of regressors, and identify the most relevant factors in explaining the success of fiscal con-

[3]Romer and Romer (2010) identify large tax policy actions in the US according to the narrative record, such as presidential speeches and Congressional reports. Ramey and Shapiro (1998) construct a war dummy that identifies large increases in government expenditure due to military reasons in the US.

[4]There is an enormous VAR literature analysing the macroeconomic effects of fiscal policy. Methods for separately identifying government expenditure and government revenue shocks in VARs have been developed in the work of Blanchard and Perotti (2002) and Mountford and Uhlig (2009). Within the VAR framework, Perotti (1999) finds that there is a higher probability of fiscal policy being non-Keynesian when there is a significantly high public debt-to-GDP ratio.

solidation programs. Empirical results indicate that favorable economic conditions in terms of small output gaps are found to be the only relevant factor that unambiguously generates successful (in terms of reducing debt-to-GDP ratios) fiscal consolidations. If the focus is on the persistence of the primary deficit reductions, cutting public wages seems to be the only crucial factor explaining successful fiscal adjustments.[5]

The remainder of the paper is organised as follows. The data used in the paper are described in Section 2. Section 3 illustrates the potential endogeneity of fiscal consolidation episodes, and Section 4 describes the empirical approach considered to estimate causal effects from fiscal consolidations to economic performance. In Section 5 we present the empirical results. Section 6 investigates the main determinants of successful consolidations. Finally, some concluding remarks are discussed in Section 7.

# 2 Data

The data used in this paper are from the OECD Economic Outlook No. 84. The sample includes annual information for 21 OECD countries (Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany,[6] Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and the United States) from 1980 to 2007.

The fiscal variables included in the dataset can be classified in three different groups: (i) those variables related to the fiscal situation of the country (e.g. government debt and primary budget balance as a share of GDP); (ii) variables capturing the composition of the fiscal consolidation (or stimulus) program (e.g. current primary expenditure, government wage and non-wage expenditures, subsidies, income taxes, social security contributions...); and finally, (iii) the change in cyclically adjusted primary deficit as a proxy for the size of the fiscal manoeuvre. The dataset also incorporates a set of macro variables such as the GDP growth rate, the exchange rate, and the short-term interest rate.

Finally, we also consider in the dataset two different dummy variables identifying the consolidation episodes. On the one hand, we have the dummy from Alesina and Ardagna (2010) — henceforth AA2010 — that identifies a fiscal consolidation episode in a given year if the cyclically adjusted primary balance (CAPB) improves by at least 1.5 per cent of GDP (analogously, a stimulus dummy takes the value 1 if the CAPB deteriorates by at least 1.5 per cent of GDP). Alternatively, we also consider the IMF consolidation dummy[7] defined following the narrative approach and focusing on policy actions (i.e. years in which the government implemented tax hikes or spending cuts to reduce the budget deficit regardless

[5]The rationale for this result comes from the investment channel described in Alesina et al. (2002) who emphasize that deficit reductions achieved through spending cuts from the wage bill (rather than tax increases) are more likely to be successful. This is so because cutting public wages might generate downward wage pressures in the private sector that result in higher levels of investment.

[6]The data for Germany starts in 1992.

[7]The IMF dummy is only available for 15 countries instead of 21. In particular, this variable is not available for Austria, Greece, Netherlands, New Zealand, Norway and Switzerland.

of the change in the CAPB).

Additional information on the variables considered can be found in the Appendix.

# 3 Are Fiscal Consolidation Episodes Endogenous?

In this section we provide some motivating and heuristic evidence of our main concern throughout the paper, the potential endogeneity of fiscal consolidation episodes. Furthermore, in Section 5 we discuss additional econometric tests which give support to the idea that fiscal consolidations are not exogenous to GDP growth.

It is reasonable to argue that the decision to make a fiscal adjustment by the fiscal authorities is not exogenous to developments in the economy. For example, agents might now anticipate fiscal retrenchments in view of recent increases in budget deficits as a consequence of support packages. The deficit consolidation decision is not an exogenous and unanticipated shock to economic agents, an essential prerequisite to give a causal interpretation to previous non-Keynesian findings in the literature.

In order to informally test this intuition we run country-by-country probits of the consolidation dummy on two lags of GDP growth and primary budget deficits. We then estimate the probabilities of a fiscal consolidation being carried out in a given year based on past economic outcomes. These predicted probabilities might be interpreted as the agents' expectations of fiscal adjustments given the course of the economy. In Figure 1 we plot the real consolidation episodes according to the AA2010 definition (top panel) and the IMF definition (bottom panel) represented as grey areas, and a solid line corresponding to the predicted probabilities (or agents' expectations) for a group of six OECD countries (Italy, United Kingdom, Denmark, Japan, Ireland and Australia).

[Figure 1 here]

Looking at the top panel of Figure 1, the AA2010 dummy seems to be fairly endogenous in the sense that it is very well predicted from past economic outcomes so that agents can easily anticipate the consolidation episode according to this definition. We argue that this result invalidates the exogeneity assumption of fiscal consolidation episodes implicitly considered in earlier literature.

On the other hand, although some consolidation episodes defined according to the IMF narrative criterion in the bottom panel are well predicted from past economic information, some others are not predicted at all. One interpretation is that, as expected from the narrative approach, the IMF dummy is "less endogenous" than the AA2010 dummy. Still, there seems to be an endogeneity problem also with this definition.

In our view, Figure 1 is illustrative of the fiscal consolidation endogeneity we address in this paper. Moreover, Section 5 provides formal evidence on the potential endogeneity of fiscal consolidation to GDP growth.

# 4 Empirical Approach: Endogeneity vs. Exogeneity

One possible approach for estimating the effect of fiscal consolidations on GDP growth is based on a regression of GDP growth between $t$ and $t-1$ ($g_t$) on a consolidation dummy ($D_t$) which takes the value one at period $t$ if there was a fiscal consolidation in this year and zero otherwise.[8] In general, studies along these lines implicitly assume that fiscal authorities ignore developments in GDP when taking the decision to make a fiscal consolidation and thus the consolidation is an unanticipated shock to agents in the economy (i.e. the consolidation dummy is exogenous to GDP). Since the evidence presented in the previous section seems to contradict the exogeneity assumption, in this paper we aim to relax it. In particular, our working assumption is that fiscal authorities, when deciding on fiscal policy[9] in year $t$, take into account developments in the economy up to this year but do not anticipate the future. Given this assumption, fiscal consolidations are no longer unanticipated shocks to economic agents, i.e., the consolidation dummy is partially endogenous to GDP.[10] Under this partially endogenous assumption, we can make use of past consolidation episodes as instruments for the current consolidation decision. Note that alternatively we might also assume that fiscal authorities can also anticipate future GDP (i.e. the consolidation decision is correlated with the full path of GDP growth, past and future). However, identification under this full endogeneity assumption, which is probably more realistic, requires extra sources of variation correlated with fiscal policy but fully uncorrelated with GDP, which are not available to the best of our knowledge. Therefore, we move from the implausible full exogeneity of fiscal consolidations to the more desirable hypothesis of partially endogenous consolidations,[11] but having in mind the possibility of fully endogenous consolidations which is beyond the scope of this paper.

Formally, the panel data model to be estimated is as follows:

$$g_{it} = \alpha g_{it-1} + \beta D_{it} + \eta_i + \delta_t + v_{it} \tag{1}$$

where $g_{it}$ represents the GDP growth rate for country $i$ ($i = 1, ..., N$) in year $t$ ($t = 1, ..., T$), and $D_{it}$ is the consolidation indicator for the same country in the same year. The model also includes country-specific unobserved heterogeneity ($\eta_i$) as well as a set of time-varying common factors ($\delta_t$) which allow the existence of cross-sectional correlations across different countries in a given period.

In this framework, the two alternative assumptions regarding the endogeneity of the

[8] Note that this is equivalent to the comparison of GDP growth means before and after the consolidation episode commonly-used in the literature (e.g. Alesina and Ardagna, 2010).

[9] Note that fiscal policy in our framework is restricted to the binary decision of making a fiscal adjustment or not.

[10] In the panel data terminology, partially endogenous regressors are also known as predetermined or weakly exogenous regressors.

[11] For the sake of simplicity, we will refer to partially endogenous consolidations as endogenous consolidations in the remaining of the paper.

fiscal tightening decision are given by:

$$E(v_{it} \mid g_i, D_i, \eta_i, \delta_t) = 0 \qquad \text{(EXOGENEITY)} \tag{2}$$

$$E(v_{it} \mid g_i^{t-1}, D_i^t, \eta_i, \delta_t) = 0 \qquad \text{(ENDOGENEITY)} \tag{3}$$

where $g_i = (g_{i1}, ..., g_{it}, ..., g_{iT})'$, $D_i = (D_{i1}, ..., D_{it}, ..., D_{iT})'$, $g_i^{t-1} = (g_{i1}, g_{i2}, ..., g_{it-1})'$, and $D_i^t = (D_{i1}, ..., D_{it})'$.

Note that, for the purpose of our paper, the key difference between the two alternatives (namely, exogeneity and endogeneity) is given by the element $D_i$ versus $D_i^t$ in the conditioning set. In the exogeneity case, $D_i$ indicates that the full path of consolidations for a given country $i$ is independent of the shock to GDP in period $t$. In contrast, in the endogeneity assumption (3), $D_i^t$ implies that only past consolidation episodes are independent of the current shock to GDP growth while future consolidations will be affected by current GDP growth; hence we can label these consolidations as endogenous. Additionally, note that an analogous reasoning applies to the case of the lagged dependent variable ($g_{it-1}$) which, given the dynamics of the model, is endogenous by construction. Finally, correlation between the country-specific effects ($\eta_i$) and the regressors ($g_{it-1}$ and $D_{it}$) is allowed.

In order to estimate the model under the exogeneity assumption (2) we make use of a panel OLS estimator with country-specific effects. On the other hand, in order to accomodate the endogeneity assumption (3) several estimators are available in the literature. The most common approach to handling the presence of fixed effects and endogenous regressors is to first-difference the data and use the panel IV or GMM estimators suggested in Anderson and Hsiao (1982) and Arellano and Bond (1991). The intuition behind both estimators is based on using lagged levels of the variables as instruments of their first differences.[12] More concretely, Anderson and Hsiao (1982) propose to use one lag as instrument, while, in order to gain efficiency in the estimates, Arellano and Bond (1991) suggest a particular combination of all available lags as instruments.

In a panel setting in which neither $T$ is small nor $N$ is large (as it is our case with around 20 countries and 20 time periods) the proliferation of reduced form coefficients is a concern in the Arellano and Bond's (1991) estimator[13] (see Arellano, 2004). This concern is even worse in our case because the correlation between the instruments and the endogenous variables is not very high and the potential weak instrument bias is larger the higher the number of instruments considered (see Anderson et al., 1982). All in all, our preferred option is the panel IV estimator suggested by Anderson and Hsiao (1982) because it only uses one instrument for each endogenous variable and the number of reduced form coefficients does not grow with $T$.

[12] Another alternative is the use of the so-called system-GMM estimator introduced by Arellano and Bover (1995) and Blundell and Bond (1998) that also exploits first differences of the variables as instruments for the equation in levels. However, this estimator requires the additional assumption of mean stationarity of the variables, which in the case of GDP growth is very restrictive.

[13] Intuitively, with $N = 20$ and $T = 20$, some reduced form equations would be linear projections with $N = 20$ observations and $T - 1 = 19$ regressors.

The implementation of the Anderson and Hsiao (1982) estimator consists of a regression of $\Delta g_{it}$ on $\Delta g_{it-1}$ and $\Delta D_{it}$ using $g_{it-2}$ and $D_{it-1}$ as intruments. More concretely, the moment conditions implied by the Anderson and Hsiao's (1982) estimator —henceforth panel IV— can be expressed as follows:

$$E(\sum_{t=2}^{T} g_{it-2}\Delta v_{it}) = 0 \quad (4)$$

$$E(\sum_{t=1}^{T} D_{it-1}\Delta v_{it}) = 0 \quad (5)$$

The validity of the instruments considered in (4) and (5) relies on the validity of the endogeneity assumption in (3) which can be tested. Note that the endogeneity assumption (3) implies lack of autocorrelation in $v_{it}$ since lagged $v$s are linear combinations of the variables in the conditioning set. Moreover, if the errors in levels ($v$s) are serially independent, those in first-differences will exhibit first- but not second- order serial correlation. Therefore, testing for lack of second-order autocorrelation in $\Delta v_{it}$ is equivalent to testing the validity of assumption (3) and thus the validity of the instruments considered in our panel IV approach. For this purpose, we will make use of the AR(2) test suggested by Arellano and Bond (1991).

Finally, under the exogeneity assumption in (2), both panel OLS and panel IV estimators are consistent but panel OLS is more efficient. However, under the endogeneity assumption in (3), panel IV is consistent while panel OLS is not. Therefore, the validity of the exogeneity assumption against the endogeneity alternative will also be tested by means of a Hausman test in the next section.

This paper presents the different estimates obtained under the two alternative assumptions. Anticipating the results, while the estimates based on exogeneity produce the so-called non-Keynesian effects (i.e. $\beta > 0$), the estimates under endogeneity do not (i.e. $\beta < 0$). The results obtained from the version under exogeneity are in line with previous literature. However, by relaxing the exogeneity assumption and allowing a very simple version of endogeneity, we obtain the expected effects in a Keynesian framework (i.e. fiscal consolidations negatively affect GDP growth in the short run). This is true for the consolidation dummy defined as in AA2010. For the IMF definition we obtain that in both cases the effect is negative, confirming the evidence in the previous section that the IMF dummy is "less endogenous". Moreover, while Hausman test results provide evidence against the validity of the exogeneity assumption, AR(2) tests provide evidence in favor of the validity of our panel IV approach.

# 5 Results

Table 1 presents the results from estimating the model in (1) under both the exogeneity assumption in (2) and the endogeneity assumption in (3). In particular we regress GDP

growth on lagged GDP growth and a fiscal consolidation[14] dummy defined according to either the AA2010 CAPB-based criterion or the IMF narrative approach. The estimation is conducted using a panel OLS estimator for the exogenous consolidation case and the panel IV approach previously described for the endogeneity case.

[Table 1 here]

Columns (1) and (2) of Table 1 present the results for the CAPB-based definition of fiscal consolidation considered by AA2010. More concretely, assuming exogeneity of the consolidation in column (1), the coefficient on the consolidation dummy is positive and significant. Therefore we replicate the non-Keynesian result previously found in the literature that the fiscal multiplier might be negative (e.g. Giavazzi and Pagano, 1990; Alesina and Ardagna, 1998). However, the sharp contrast once we account for the possible endogeneity of the consolidation decision in column (2) is enlightening. In particular, taking into account the feedback effects from GDP to fiscal policy decisions, we recover the typical Keynesian result and find that the fiscal adjustment produces a decrease in GDP in the short term (i.e. the coefficient on the fiscal consolidation dummy is now negative and significant). Moreover, the AR(2) test points to the validity of the instruments considered in the panel IV approach because the hypothesis of lack of second order autocorrelation in the first-differenced errors cannot be rejected (note that this hypothesis is a consequence of the endogeneity assumption in (3) as explained in the previous section). Finally, the Hausman test provides clear evidence against the validity of the exogeneity assumption commonly considered in the literature, which reinforces the heuristic evidence presented in Section 3.

Turning to the IMF definition of fiscal consolidation in columns (3) and (4), we find that the impact of the fiscal retrenchment episode is negative in both cases (i.e. under exogeneity by panel OLS and under endogeneity by panel IV). However, while the negative coefficient under exogeneity is not significant at conventional levels, the estimate under endogeneity is more negative and highly significant. Thus we conclude that, in the short run, IMF narrative consolidations also have a negative effect on GDP growth as expected in a Keynesian framework.[15] The AR(2) test provides evidence in favor of the validity of the instruments considered in our panel IV estimates and thus in favor of the endogeneity assumption in (3). Finally, according to the Hausman test, the exogeneity of the IMF dummy cannot be rejected at the 5% significance level but can be rejected at the 10%. This evidence is again in line with the motivational evidence provided in Figure 1: fiscal consolidations defined according to the narrative approach by the IMF are, to some extent, exogenous to GDP growth.

Finally, columns (5) and (6) present the results for a fiscal stimulus dummy defined according to the CAPB-based critetion by AA2010. In this case we observe the same pattern, while stimuli assumed to be exogenous (column (5)) produce a negative and significant effect

[14] We also present the results considering the fiscal stimulus dummy defined by AA2010.

[15] The associated fiscal multipliers under endogeneity of the fiscal consolidation range from 0.2 using the AA2010 definition to 0.5 considering the IMF definition.

on output (i.e. non-Keynesian effect), under the assumption of endogenous stimuli in column (6) this non-Keynesian effect vanishes. However, in contrast to the consolidation cases, while the AR(2) test points to the validity of assumption (3), the Hausman test result indicates that exogeneity cannot be rejected.

## 5.1 Heterogeneous Effects

According to Blanchard (1990) and Perotti (1999), if consolidation is undertaken starting from a low level of current debt, a traditional positive fiscal multiplier will ensue. If, instead, fiscal consolidation is made starting from a high debt level, non-Keynesian effects via consumption might appear as a result of an expected increase in permanent income. The reason for this is that by consolidating now, the government will not raise taxes too much in the future to pay back the debt. This reduces the dead-weight loss imposed by taxes, thus raising agents' permanent income.

Alternatively, non-Keynesian effects of consolidations may take place via the behavior of investment. The link between fiscal policy and investment behavior is formalized in Alesina et al. (2002). The main effect is represented by the impact of the government wage bill on the labor market. Investment decisions by firms are driven by the expected present value of the net marginal product of capital, which in turn is a negative function of real wages. Fiscal consolidations obtained through expenditure cuts can then reduce wage pressures and so increase short-run investments. This hypothesis crucially depends upon the composition of adjustment (expenditure cuts, particularly the wage bill, versus tax increases) and on institutional factors such as the functioning of the labor market.

On the other hand, according to the standard Keynesian view, a fiscal consolidation might be expansionary (i.e. it might be followed by revived economic growth) if it is accompanied by a sufficiently lax monetary policy. Therefore a reduction in the interest rate, or a devaluation in the case of a small economy, might generate economic growth during the process of a fiscal consolidation.

Finally, during the year 2010 financial crises have exerted an important effect on fiscal policy and in particular on the timing of fiscal consolidation programs. If fiscal retrenchments are clearly necessary (for instance because the country is paying unsustainable prices for its debt), once the government finally undertakes the consolidation, economic agents might react optimistically increasing consumption and/or investment because they have already discounted the adjustment.

In order to further investigate the existence of expansionary effects through the channels discussed above, we consider the following specification:

$$g_{it} = \alpha g_{it-1} + \beta D_{it} + \gamma Z_{it} D_{it} + \eta_i + \delta_t + v_{it} \tag{6}$$

where $Z_{it}$ is one of the following variables: the government level of debt, the proportion of the adjustment due to cuts in current expenditures and more particularly in the public

wage bill, a structural reform dummy[16] capturing changes in institutional factors during the fiscal consolidation episode, changes in the interest rate and the exchange rate to capture monetary policy movements, a dummy for those country-year under the Maastricht Treaty,[17] and finally, the spread of the ten-year government bond with respect to Japan as a proxy of financial turmoil in the country. These eight variables aim to explain the possibility of fiscal consolidation followed by economic growth.

Note that now the effect of the consolidation on GDP is given by:

$$\phi = \beta + \gamma Z_{it} \tag{7}$$

so that even if $\beta$ is negative, $\phi$ might become positive for certain values of the $Z$ variables; for instance, for a sufficiently high level of debt or interest rate spread, for a sufficiently high cut in public wages, for those consolidations accompanied by pro-market structural reforms or expansive monetary policy.

Table 2 presents the results for estimating equation (6) for the eight different $Z$ variables considered using the two alternative consolidation dummies available (CAPB-based from AA2010, and action-based from the IMF). As described in the previous section, we estimate the equation under endogeneity.

[Table 2 here]

The main conclusion from the estimates in Table 2 is that, in general, there is no positive effect from fiscal consolidations to GDP growth through the channels tested (i.e. $\phi$ is negative). Neither the investment channel nor the consumption channel of non-Keynesian effects seem to be at work according to these results (see columns from (1) to (8)). On the other hand, neither the Masstricht dummy nor the spread of the ten-year government bonds seem to explain consolidations followed by economic growth (see columns (11), (12), (15), and (16) in Table 2).

However, the estimates in columns (9) and (13) indicate that a positive coefficient arises when the fiscal consolidation is accompanied by lax monetary policy.[18] In particular, if the fiscal adjustment is carried out together with a reduction in the short-term interest

[16] We employ the reform dummy in Duval (2008). For the construction of this dummy, we first consider an overall index of rigidity measuring the anti-competitive/distortionary effects of policies for the 21 OECD countries included in our sample over the period 1985-2003 in five policy areas: labor taxes, unemployment benefit system, employment protection legislation (EPL), retirement schemes and product market regulations. Secondly, given this index, a reform in a given policy field is identified by a dummy taking value 1 whenever the corresponding index of rigidity falls sufficiently. As a benchmark, the requirement is a change in the index to be below the 20th percentile of its distribution across the whole sample. A labor market reform dummy from Boeri and Garibaldi (2009) was also considered with the same results.

[17] In order to maintain the price stability within the Eurozone, the Maastricht criteria impose some requisites on the inflation rate, the annual government deficit, the level of government debt, and the long-term interest rate of the country members. The effects of a fiscal adjustment in countries subject to the Maastricht Treaty might differ from those in countries without these requisites.

[18] In the form of interest rate reductions. Changes in the exchange rate do not seem to account for expansionary fiscal consolidations according to the results in columns (10) and (14).

rate larger than 28%, we might expect revived economic growth in the subsequent periods. Our interpretation is that the Keynesian effects of such a monetary policy compensate the negative (and also Keynesian) effects of the fiscal retrechment.

## 5.2 Effect on Consumption, Investment, and Unemployment

In order to further investigate the channels through which fiscal consolidations affect GDP, we now turn to the estimation of the effects of fiscal consolidations on private consumption, private investment and the unemployment rate for a panel of OECD countries. van Aarle and Garretsen (2003) and Hogan (2004), among others, estimate consumption functions and test whether the impact of government revenues and expenditures on consumption is different when fiscal consolidations take place. Alesina et al. (2002) conduct the same kind of analysis but estimating investment equations instead of consumption equations. We may also interpret the estimate of the consolidation effects on consumption and/or investment as an additional test of the two non-Keynesian hypotheses in the previous section.

We again estimate equation (1) but replacing GDP growth with consumption, investment, and unemployment growth:

$$g^h_{it} = \alpha g^h_{it-1} + \beta D_{it} + \eta_i + \delta_t + v_{it} \tag{8}$$

where $g_{it}$ represents the growth rate between $t-1$ and $t$ in country $i$, $h$=(consumption, investment, unemployment), and the remaining terms are as in equation (1). We estimate this model under endogeneity of the fiscal adjustment for the two alternative definitions of consolidation (AA2010 and IMF).

[Table 3 here]

Table 3 presents the results. We observe in columns (1) and (2) that the effect of a fiscal consolidation episode on the investment growth rate is negative and significant under endogeneity of the fiscal tightening decision. Moreover, the magnitude of this effect on private investment is much larger than the effect on GDP growth. In particular, a fiscal consolidation causes, on average, a decrease in the rate of growth of investment of 55-105 basis points.

Regarding the effect on private consumption in columns (3) and (4), the evidence is mixed: while the effect of a fiscal consolidation on the consumption growth rate is negative and significant when considering the IMF action-based denition of consolidation, the estimated effect is not significant according to the CAPB-based definition by AA2010. The opposite pattern arises for the rate of growth of unemployment in columns (5) and (6). The effect is positive and significant according to the AA2010 consolidation definition but it becomes insignificant when considering the IMF definition.

# 6 Determinants of Successful Consolidations

According to AA2010, successful fiscal adjustments are those in which the cumulative reduction of the debt-to-GDP ratio three years after the beginning of the adjustment is greater than 4.5 percentage points.

What are the main characteristics of these successful fiscal consolidations? Understanding under which circumstances a fiscal consolidation might succeed in reducing the level of debt and the primary deficit is crucial from a policy perspective. In order to answer this question, the typical approach in the literature is to construct a successful consolidation dummy that takes the value one for those country-years in which the consolidation succeeded in terms of the previous definition, and zero otherwise. Then, they run a regression of this dummy on a set of macro and fiscal variables capturing the environment in which the successful consolidation took place. Depending on the $t$-statistics of this regression, they conclude which are the most relevant characteristics surrounding a successful fiscal consolidation.

In broad terms, researchers aim to disentangle the importance of four competing explanations of successful consolidations proposed in the literature: (i) the country's fiscal situation prior to the consolidation, proxied by, for example, the government debt as a share of GDP in the previous year (e.g. Perotti, 1999); (ii) the size of the adjustment proxied by the change in primary deficit during the episode (e.g. Giavazzi and Pagano, 1996; Giavazzi et al., 2000); (iii) the composition of the adjustment in terms of the change in the different items of the public bill as a share of the whole change in the primary deficit (e.g. Alesina and Perotti, 1995; McDermott and Wescott, 1996); (iv) the macroeconomic situation captured through the output gap or the growth rate of GDP (e.g. Lambertini and Tavares, 2003).

The papers by Ardagna (2004) — henceforth A04 —, Alesina and Ardagna (1998) — henceforth AA98 —, and Giudice et al. (2007) — henceforth G07 — are good examples of this approach. All the three papers regress the successful consolidation dummy on a set of regressors aiming to capture some of the four hypothesis described above. However, there is no agreement on which regressors / hypothesis must be included in the empirical model (model uncertainty), so that each of the three papers considers a different model. Table 4 presents the results we obtain replicating the regressions in these papers but using our dataset.[19]

[Table 4 here]

In view of the results presented in the AA98 column in Table 4, one might conclude that the composition in terms of wage expenditures does not affect the success of the consolidation program. However, according to the A04 column, a consolidation based on cutting

[19]Note here that some of the results are not exactly replicated because either the sample period or the variables' definition is not equal to the original papers. Nevertheless, these results are only an illustration of the model uncertainty problem for the sake of motivation.

public wages is expected to reduce the probability of success in terms of debt reduction.[20] On the other hand, the macroeconomic environment proxied by the output gap is not significantly correlated with the successful consolidation dummy according to the G07 specification. Nevertheless, if we proxy the macro conditions using the GDP growth rate, as in the A04 column, they are found to positively affect the probability of success in reducing the debt-to-GDP ratio. Moreover, these results might again change if we consider different proxies and the number of them we include in the regression (for instance, we can also include other items of the public bill such as non-wage expenditures, transfers, business taxes, income taxes...). Since the number of potential proxies for the four candidate theories (i.e. fiscal situation, consolidation size, consolidation composition, and macroeconomic situation) is enormous, the universe of potential regressions to estimate given all the possible combinations of proxies is very difficult to work with. Therefore, extracting conclusions robust to the particular regression estimated might be extremely difficult in this setting in view of the simple examples presented in Table 4.

To overcome these issues we consider model averaging methods. Model averaging represents an agnostic alternative to the previous approach (i.e. the approach considered in Table 4 which is based on selecting a single regression and deciding which variable is important depending on its associated $t$-ratio). The key idea of model averaging is to consider and estimate all the possible regressions, and then report a weighted average as the estimate of interest. Therefore, model averaging is agnostic in the sense that a researcher relying on this approach holds the view that the true single model is unknown and probably unknowable. The best she can do, then, is to consider all the possible alternatives instead of basing her conclusions on one single regression.

The model averaging methodology allows a ranking to be constructed of the variables ordered by their relative importance in the contribution to the model fit, i.e. the Posterior Inclusion Probability (PIP). Those variables with higher PIP are the ones that contribute most to explaining the dependent variable's variation, in our case the successful consolidation episodes. Note that the results presented in this paper using model averaging techniques must be interpreted with caution, we can only conclude which are the variables most robustly correlated with successful consolidations. A brief introduction and formal details of model averaging techniques can be found in the Appendix.

[Table 5 here]

Table 5 presents the results when applying model averaging to estimate all the candidate regressions in order to investigate which regressors are robust determinants of successful consolidations. Output gap is the only variable for which the posterior inclusion probability (PIP) is higher than the prior inclusion probability, whereby the main conclusion emerging from Table 5 is that the output gap is the only robust determinant of successful fiscal

[20] Ardagna (2004) concludes that stabilizations implemented by cutting public spending lead to higher GDP growth rates, and also that the success of fiscal adjustments in reducing debt-to-GDP ratio depends on the size of the contraction and less on its composition.

consolidations. Moreover, its posterior standard error is smaller than its posterior mean. Therefore we can also conclude that the output gap positively affects the probability of success of a fiscal consolidation package. This result implies that whatever the composition or the size of the adjustment, the most relevant economic policies in times of fiscal consolidation must be oriented toward the objective of sustained and higher rates of GDP growth. Those consolidations not accompanied by economic reforms aimed at increasing employment and productivity will have more difficulties in the reduction of debt-to-GDP ratios. One interpretation of this result is that it is easier to succeed by increasing the denominator (GDP) than by reducing the numerator (debt). Which are the best policies for increasing GDP is of course a controversial question that is beyond the scope of this paper.

## 6.1 Alternative Definitions of Successful Consolidations

How to define a successful consolidation is not straightforward, and the literature has considered different criteria. Once we have identified the fiscal consolidation episodes in the OECD according to the CAPB-based definition in AA2010,[21] in this sub-section we isolate successful consolidations considering two alternative criteria, in addition to the debt-to-GDP ratio criterion considered in the previous section (which is usually the most common approach in the literature).

First, we use an expansionary criterion in terms of GDP growth to identify successful consolidations. According to this expansionary criterion, a fiscal consolidation succeeds if average trend growth between $t$ and $t+2$ is greater than between $t-1$ and $t-2$ (see Giudice et al., 2007). These fiscal consolidations are usually labeled in the literature as expansionary consolidations.

Second, we also consider a persistence criterion. This criterion identifies as successful those consolidations in which the primary cyclically adjusted budget balance improves by at least three percentage points of GDP over three consecutive years (i.e. between $t-2$ and $t$, between $t-1$ and $t+1$ or between $t$ and $t+2$), and in each year the change in the primary cyclically adjusted budget balance cannot be below $-0.5$ percentage points of GDP.

Using these two dummy variables we repeat the analysis carried out in the previous section using model averaging. Results are presented in Table 6; columns (1), (2), and (3) report the robust determinants of successful consolidations defined according to the persistence criterion. With respect to expansionary consolidations, results are reported in columns (4), (5), and (6) of Table 6.

[Table 6 here]

Regarding the persistence criterion, only the change in wage expenditures as a share of the total change in the primary deficit is a robust determinant of successful consolidations. This implies that the higher the proportion of the consolidation conducted via reducing

[21] The consolidation dummy from AA2010 identifies a fiscal consolidation episode in a given year if the cyclically adjusted primary balance (CAPB) improves by at least 1.5% of GDP.

public wages, the higher the probability of the adjustment being successful in terms of persistence in the deficit reduction. Cutting public wages is a very costly political decision and, therefore, governments will only reduce the public wage bill when they take seriously the fiscal consolidation program, and they are thus more likely to achieve the deficit reduction objective.

With respect to the expansionary criterion, given that the prior inclusion probability for each variable is 0.5 and all the PIPs are below this threshold, we conclude that there is no variable robustly affecting the probability of a fiscal consolidation being expansionary. This is the most-commonly used criterion for labeling variables as robust / non-robust in the model averaging literature. We assume a priori that all variables are equally robust (i.e. prior inclusion probability of 0.5) and we label as robust those variables for which the PIP is higher than 0.5. This would imply that the data support these variables more than the rest of the regressors, but if no variable satisfies this criterion, the conclusion is that no variable is robust. In addition, all the variables considered as candidate determinants have posterior standard errors larger than the corresponding posterior means, which reinforces the previous conclusion of no variable robustly correlated with expansionary fiscal consolidations.

# 7 Concluding Remarks

The decision to carry out a fiscal adjustment in order to restore the budget balance is not independent of developments in the economy. This paper provides evidence in favor of this hypothesis and estimates the effect of fiscal retrenchments on GDP growth accounting for the endogeneity of these episodes. In particular, while under exogeneity we find that fiscal consolidations positively affect GDP growth, considering endogenous consolidations we find that fiscal adjustments have the expected negative (and Keynesian) effect on GDP growth in the short term. This is true for both the CAPB-based definition of consolidation considered in Alesina and Ardagna (2010) and the fiscal consolidations defined by the IMF (2010) following the narrative approach. This sharp contrast between the results under exogeneity and under endogeneity indicates that non-Keynesian effects of fiscal consolidation episodes previously found in the literature might be due to endogeneity biases.

Successful consolidations are those in which the reduction of the debt-to-GDP ratio three years after the beginning of the adjustment is greater than 4.5%. We analyse which factors are the most relevant in generating successful consolidations via model averaging techniques. Our results indicate that, in order to succeed in reducing budget deficits, economic growth is the only relevant ingredient. Without economic recovery, fiscal consolidations will have huge difficulties in reducing budget deficits. However, we also find that cuts in public wages are the only ingredient of fiscal consolidations in which persistent reductions in primary budget deficits were achieved.

# A Appendix

## A.1 Data Appendix

This section describes the data employed in the paper. All data are from the OECD Economic Outlook Database no. 84.

- Government debt level: government gross debt as a share of GDP.
- Deficit level: cyclically adjusted primary deficit as a share of GDP (i.e. primary expenses minus total revenue)
- Consolidation size: Change in the cyclically adjusted primary balance as a share of GDP.
- $\Delta$Wage expenditures: Change in government wage bill expenditures.
- $\Delta$Non-wage expenditures: Change in government non wage bill expenditures.
- $\Delta$Subsidies: Change in subsidies to firms.
- $\Delta$Transfers: Change in cyclically adjusted transfers as a share of GDP.
- $\Delta$Government investment: Change in the gross government consumption on fixed capital.
- $\Delta$Income taxes: Change in cyclically adjusted direct taxes on household as a share of GDP.
- $\Delta$Business taxes: Change in cyclically adjusted direct taxes on businesses as a share of GDP.
- $\Delta$Indirect taxes: Change in cyclically adjusted indirect taxes as a share of GDP.
- $\Delta$Other taxes: Change in cyclically adjusted other taxes (different from income, business or indirect) as a share of GDP.
- $\Delta$S.s. contributions: Change in cyclically adjusted social security contributions paid by employers and employees as a share of GDP.
- GDP growth: Yearly growth rate of real per capita GDP for each country.
- Output gap: % of potential GDP.
- $\Delta$Interest rate: Change in the real short-run interest rates between $t+1$ and $t-1$.
- $\Delta$Exchange rate: Change in the exchange rate between $t+1$ and $t-1$.
- Investment growth: Yearly growth rate of private non-residential gross fixed capital formation.

- Consumption growth: Yearly growth rate of private final consumption expenditure.
- Unemployment growth: Yearly growth rate of the unemployment rate.

Note that all the regressors belonging to the public bill (e.g. ΔWage expenditures, ΔSubsidies, ΔIndirect taxes,...) are divided by the total change in the primary deficit (ΔItem/ΔDeficit) to focus on the proportion of the adjustment which was due to a particular item as proxies of the composition. More concretely, an increase in these variables means that a larger share of the change in the primary deficit is due to a change in the particular item of the public bill. For the spending items, an increase in these variables means that a larger share of the increase (reduction) of the primary deficit is obtained by increasing (cutting) the particular spending item. For the revenue items, an increase in these variables means that a larger share of the increase (reduction) of the primary deficit is obtained by cutting (increasing) the particular revenue item of the government budget.

## A.2 (Bayesian) Model Averaging

Within the paper's framework, model uncertainty arises because the lack of clear theoretical guidance on the choice of regressors affecting the probability of success during fiscal consolidation episodes results in a wide set of possible specifications. Therefore, researchers' uncertainty about the value of the coefficient of interest in a regression exists at two distinct levels. The first is the uncertainty associated with the parameter conditional on a given empirical model. This level of uncertainty is of course assessed in virtually every empirical study. What is not fully assessed is the uncertainty associated with the specification of the empirical model. It is typical for a given paper that the specification of the regression is taken as essentially known; while some variations of a baseline model are often reported, via different choices of control variables, standard empirical practice does not systematically account for the sensitivity of claims about the parameter of interest to model choice.

Many researchers consider that the most promising approach to account for model uncertainty is to employ model averaging techniques to construct parameter estimates that formally address the dependence of model-specific estimates on a given model. The basic idea behind model averaging is to estimate the distribution of unknown parameters of interest across different models. The fundamental principle of model averaging is to treat models and related parameters as unobservable, and to estimate their distributions based on the observable data. In contrast to classical estimation, model averaging copes with model uncertainty by allowing for all possible models to be considered, which consequently reduces the biases of parameters.

Formally, consider a generic representation of an empirical model of the form:

$$\Psi = \theta X + \epsilon \tag{9}$$

where $\Psi$ is the dependent variable of interest (the successful consolidation dummy in our case), and $X$ represents a set of covariates (such as the level of government debt, the size of

the consolidation...). Imagine that there are potentially very many empirical models, each given by a different combination of explanatory variables (i.e. different vectors $X$), and each with some probability of being the 'true' model. This is the starting idea of the Bayesian Model Averaging methodology.[22]

Using the Bayesian jargon, a model is formally defined by a likelihood function and a prior density. Suppose we have $K$ possible explanatory variables. We will have $2^K$ possible combinations of regressors, that is to say, $2^K$ different models - indexed by $M_j$ for $j = 1, ..., 2^K$- which all seek to explain $y$ -the data-. $M_j$ depends upon parameters $\theta^j$. In cases where many models are being entertained, it is important to be explicit about which model is under consideration. Hence, the posterior for the parameters calculated using $M_j$ is written as:

$$g\left(\theta^j|y, M_j\right) = \frac{f\left(y|\theta^j, M_j\right) g\left(\theta^j|M_j\right)}{f\left(y|M_j\right)} \tag{10}$$

and the notation makes clear that we now have a posterior, a likelihood, and a prior for each model. The logic of Bayesian inference suggests that we use Bayes' rule to derive a probability statement about what we do not know (*i.e.* whether a model is correct or not) conditional on what we do know (*i.e.* the data). This means the posterior model probability can be used to assess the degree of support for $M_j$. Given the prior model probability $P\left(M_j\right)$ we can calculate the posterior model probability using Bayes Rule as:

$$P\left(M_j|y\right) = \frac{f\left(y|M_j\right) P\left(M_j\right)}{f\left(y\right)} \tag{11}$$

Since $P\left(M_j\right)$ does not involve the data, it measures how likely we believe $M_j$ to be the correct model before seeing the data. $f\left(y|M_j\right)$ is often called the marginal (or integrated) likelihood, and is calculated using (10) and a few simple manipulations. In particular, if we integrate both sides of (10) with respect to $\theta^j$, use the fact that $\int g\left(\theta^j|y, M_j\right) d\theta^j = 1$ (since probability density functions integrate to one), and rearrange, we obtain:

$$f\left(y|M_j\right) = \int f\left(y|\theta^j, M_j\right) g\left(\theta^j|M_j\right) d\theta^j \tag{12}$$

The quantity $f\left(y|M_j\right)$ given by equation (12) is the marginal probability of the data, because it is obtained by integrating the joint density of $(y, \theta^j)$ given $y$ over $\theta^j$. The ratio of integrated likelihoods of two different models is the Bayes Factor and it is closely related to the likelihood ratio statistic, in which the parameters $\theta^j$ are eliminated by maximization rather than by integration.

Moreover, considering $\theta$ a function of $\theta^j$ for each $j = 1, ..., 2^K$, we can also calculate the posterior density of the parameters for all the models under consideration:

$$g\left(\theta|y\right) = \sum\nolimits_{j=1}^{2^K} P\left(M_j|y\right) g\left(\theta|y, M_j\right) \tag{13}$$

[22]While model averaging can be interpreted from a frequentist viewpoint, its roots are based on the Bayesian paradigm. See Moral-Benito (2010) for an overview of model averaging methods.

If one is interested in point estimates of the parameters, one common procedure is to take expectations across (13):

$$E\left(\theta|y\right) = \sum\nolimits_{j=1}^{2^K} P\left(M_j|y\right) E\left(\theta|y, M_j\right) \tag{14}$$

Following Leamer (1978), we calculate the posterior variance as:

$$\begin{aligned} V\left(\theta|y\right) &= \sum\nolimits_{j=1}^{2^K} P\left(M_j|y\right) V\left(\theta|y, M_j\right) \\ &+ \sum\nolimits_{j=1}^{2^K} P\left(M_j|y\right) \left(E\left(\theta|y, M_j\right) - E\left(\theta|y\right)\right)^2 \end{aligned} \tag{15}$$

Inspection of (15) shows that the posterior variance incorporates both the estimated variances of the individual models as well as the variance in estimates of the $\theta$'s across different models. Hence, the uncertainty at the two different levels mentioned above is taken into account.

Moreover, the BMA methodology allows constructing a ranking of variables ordered by their robustness. In our particular case, robustness as determinants of successful fiscal consolidations. In order to construct our measure of robustness, we estimate the posterior probability that a particular variable $h$ is included in the regression, and we interpret it as the probability of that the variable belongs in the true empirical model. In other words, variables with high posterior probabilities of being included are considered as *robust* determinants of succeed when a fiscal adjustment is carried out. This is called the *posterior inclusion probability* for variable $h$, and it is calculated as the sum of the posterior model probabilities for all of the models including that variable:

$$\text{posterior inclusion probability} = P\left(\theta_h \neq 0|y\right) = \sum\nolimits_{\theta_h \neq 0} P\left(M_j|y\right) \tag{16}$$

As an indication of our ignorance, we assume that all the possible models are equally probable a priori so that $P(M_j = 1/2 \quad \forall \quad j = 1, ..., 2^K$. This prior on the model space also implies a prior on the regressors, in particular, it implies that all regressors have a prior inclusion probability equal to 0.5. It is usual in the model averaging literature to impose a threshold to determine which variables are robust. More concretely, the most commonly-used threshold is the prior inclusion probability, i.e. those regressors with posterior inclusion probability higher than the prior inclusion probability are labeled as robust because the data supports their inclusion in the model.

On the other hand, we make use of the Schwarz asymptotic approximation to the Bayes Factor, and therefore replace equation (11) by:

$$P(M_j|y) = \frac{P(M_j) N^{\frac{-k_j}{2}} SSE_j^{-\frac{N}{2}}}{\sum_{i=1}^{2^K} P\left(M_i\right) N^{\frac{-k_i}{2}} SSE_j^{-\frac{N}{2}}} \tag{17}$$

where $SSE_j$ is the sum of squares for model $j$, and $N$ is the number of observations. Therefore, instead of equation (14) we will use:

$$E(\theta|y) = \sum_{j=1}^{2^K} P(M_j|y) E(\theta|y, M_j) = \sum_{j=1}^{2^K} P(M_j|y) \widehat{\theta}^j_{OLS} \tag{18}$$

where $\widehat{\theta}^j_{OLS}$ is the OLS estimate for model $j$. Equation (18) is true if we either assume diffuse priors on the parameter space for any given sample size, or have a large sample for any given prior on the parameter space. Equations (17) and (18) are the basis of the BACE approach described in Sala-i-Martin et al. (2004) in the context of growth regressions.

# References


[1] Alesina, Alberto, and Silvia Ardagna. 1998. "Tales of Fiscal Contractions." *Economic Policy*, 27: 487-545.

[2] Alesina, Alberto, and Silvia Ardagna. 2010. "Large Changes in Fiscal Policy: Taxes versus Spending." *Tax Policy and the Economy*, Editor: R. Brown, vol. 24. NBER.

[3] Alesina, Alberto, and Roberto Perotti (1995) "Fiscal Expansions and Adjustments in OECD countries." *Economic Policy*, 21: 205-248.

[4] Alesina, Alberto, Silvia Ardagna, Roberto Perotti, and Fabio Schiantarelli. 2002. "Fiscal policy, profits, and investment." *American Economic Review*, 92: 571-589.

[5] Anderson, Theodore, and Cheng Hsiao. 1982. "Formulation and estimation of dynamic models using panel data." *Journal of Econometrics*, 18: 47-82.

[6] Anderson, Theodore, Naoto Kunitomo, and Takamitsu Sawa. 1982. "Evaluation of the Distribution Function of the Limited Information Maximum Likelihood Estimator." *Econometrica*, 50: 1009-1027.

[7] Ardagna, Silvia. 2004. "Fiscal Stabilizations: When Do They Work and Why." *European Economic Review*, 48: 1047-1074.

[8] Arellano, Manuel. 2004. "Modelling Optimal Instrumental Variables for Dynamic Panel Data Models." mimeo.

[9] Arellano, Manuel, and Steve Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58: 277-297.

[10] Arellano, Manuel, and Olimpia Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error-Components Models." *Journal of Econometrics*, 68: 29-52.

[11] Blanchard, Olivier. 1990. "Comment on Giavazzi and Pagano." In: Blanchard O. and S. Fischer (eds) *NBER macroecomics annual*. MIT, Cambridge MA.

[12] Blanchard, Olivier. 1993. "Suggestion for a New Set of Fiscal Indicators." OECD Working paper No. 79.

[13] Blanchard, Olivier, and Roberto Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *The Quarterly Journal of Economics*, 117: 1329-1368.

[14] Blundell, Richard, and Steve Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics*, 87: 115-143.

[15] Boeri, Tito, and Pietro Garibaldi. 2009. "Beyond Eurosclerosis." *Economic Policy*, 24: 409-461.

[16] Cour, Philippine, Eric Dubois, Selma Mahfouz, and Jean Pisany-Ferry. 1996. "The cost of fiscal retrechment revisited: how strong is the evidence?" CEPII Working Paper 96-16.

[17] Duval, Romain. 2008. "Is there a role for macroeconomic policy in fostering structural reforms? Panel evidence from OECD countries over the past two decades." *European Journal of Political Economy*, 24: 491-502.

[18] Feldstein, Martin. 1982. "Government Deficits and Aggregate Demand." *Journal of Monetary Economics*, 9: 1-20.

[19] Giavazzi, Francesco, Tullio Jappelli, and Marco Pagano. 2000. "Searching for non-linear effects of fiscal policy: evidence for industrial and developing countries." *European Economic Review*, 44: 1269-1289.

[20] Giavazzi, Francesco, and Marco Pagano. 1990. "Can Severe Fiscal Contractions Be Expansionary? Tales of Two Small European Countries." *NBER Macroeconomics Annual*, pp. 95-122, MIT Press.

[21] Giavazzi, Francesco, and Marco Pagano. 1996. "Non-keynesian Effects of Fiscal Policy Changes: International Evidence and the Swedish Experience." *Swedish Economic Policy Review*, 3: 67-103.

[22] Giudice, Gabriele, Alessandro Turrini, and Jan in't Veld. 2007. "Non-Keynesian Fiscal Adjustments? A Close Look at Expansionary Fiscal Consolidations in the EU." *Open Economies Review*, 18: 613-630.

[23] Hogan, Vincent. 2004. "Expansionary fiscal contractions? Evidence from panel data." *Scandinavian Journal of Economics*, 106: 647-659.

[24] International Monetary Fund (IMF). 2010. "Will It Hurt? Macroeconomic Effects of Fiscal Consolidation." In World Economic Outlook.

[25] Lambertini, Luisa, and Jose Tavares. 2003. "Exchange Rates and Fiscal Adjustments: Evidence from the OECD and Implications for EMU." Boston College Working Paper 576.

[26] Leamer, Edward. 1978. "Specification Searches." New York: John Wiley and Sons.

[27] McDermott, John, and Robert Wescott. 1996. "An Empirical Analysis of Fiscal Adjustments." IMF Working Paper No.96/59.

[28] Miller, Stephen, and Frank Russek. 2003. "The Relationship between large fiscal adjustments and short-term output growth under alternative fiscal policy regimes." *Contemporary Economic Policy*, 21: 41-58.

[29] Moral-Benito, Enrique. 2010. "Model Averaging in Economics." CEMFI Working Paper 1008.

[30] Mountford, Andrew, and Harald Uhlig. 2009. "What Are the Effects of Fiscal Plicy Shocks?" *Journal of Applied Econometrics*, 24: 960-992.

[31] Perotti, Roberto. 1999. "Fiscal Policy In Good Times And Bad." *The Quarterly Journal of Economics*, 114: 1399-1436.

[32] Ramey, Valerie, and Matthew Shapiro. 1998. "Costly Capital Reallocation and the Effects of Government Spending." *Carnegie-Rochester Conference Series on Public Policy*, 48: 145-94.

[33] Romer, Cristina, and David Romer. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review*, 100: 763-801.

[34] Sala-i-Martin, Xavier, Gernot Doppelhofer, and Ronald Miller. 2004. "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach." *American Economic Review*, 94: 813-835.

[35] van Aarle, Bas, and Harry Garretsen. 2003. "Keynesian, non-Keynesian, or no effects of fiscal policy changes? The EMU case." *Journal of Macroeconomics*, 25: 213-240.

Table 1: Endogenous versus Exogenous Fiscal Consolidations

| | AA2010 Dummy Consolidation | | IMF Dummy Comsolidation | | AA2010 Dummy Stimulus | |
|---|---|---|---|---|---|---|
| Estimation Method | Panel OLS | Panel IV | Panel OLS | Panel IV | Panel OLS | Panel IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Dependent Variable is GDP Growth | | | | | |
| $\text{Growth}_{t-1}$ | 0.835 | 0.861 | 0.841 | 0.876 | 0.840 | 0.866 |
| | (0.020) | (0.030) | (0.023) | (0.037) | (0.019) | (0.029) |
| Consolidation / | 0.004 | −0.004 | −0.005 | −0.015 | −0.007 | −0.004 |
| Stimulus | (0.001) | (0.002) | (0.003) | (0.006) | (0.002) | (0.003) |
| Fixed Effects | YES | YES | YES | YES | YES | YES |
| AR(2) test p-value | | 0.45 | | 0.57 | | 0.31 |
| Hausman p-value | | 0.00 | | 0.07 | | 0.27 |
| Observations | 554 | 554 | 393 | 393 | 554 | 554 |
| Countries | 21 | 21 | 15 | 15 | 21 | 21 |
| R-squared | 0.70 | | 0.71 | | 0.71 | |

*Notes:* This Table presents the results of estimating equation (1) [$\text{Growth}_{it} = \alpha\ \text{Growth}_{it-1} + \beta\ \text{Consolidation}_{it} + \eta_i + \delta_t + v_{it}$] under diferent exogeneity/endogeneity assumptions and for two alternative consolidation dummies (AA2010 and IMF). In columns (5) and (6) the consolidation dummy is substituted by the stimulus dummy defined in Alesina and Ardagna (2010). Panel OLS with country-specific effects is the estimator considered in columns (1), (3) and (5) under the exogeneity assumption, and the Anderson and Hsiao (1982) panel IV estimator is employed in columns (2), (4) and (6) under the endogeneity assumption. AR(2) test p-value refers to the p-values from the Arellano and Bond's (1991) test for the lack of second order autocorrelation in the first differenced errors. Under this null hypothesis the instruments considered to address the endogeneity are valid because the errors in levels are not correlated as implied by assumption (3). Hausman p-value reports the obtained p-value from a Wald test of the null of exogenous fiscal consolidations (i.e. the null of equality between the coefficients estimated by panel OLS —efficient under the null— and panel IV —consistent under the null but robust to partial endogeneity—). If exogeneity of the fiscal consolidation holds, the null of equality of coefficients would be true. Standard errors reported in parentheses are clustered at the country level.

Table 2: Heterogeneous Effects of Fiscal Consolidations

| | **AA2010 Dummy** | | | | **IMF Dummy** | | | |
|---|---|---|---|---|---|---|---|---|
| Z Variable | Government Debt Level | Public Wages | Current Expenditure | Structural Reform | Government Debt Level | Public Wages | Current Expenditure | Structural Reform |
| Estimation Method | Panel IV | Panel IV | Panel IV | Panel IV | Panel IV | Panel IV | Panel IV | Panel IV |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\text{Growth}_{t-1}$ | 0.85 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 |
| (t-ratio) | (0.02) | (0.02) | (0.02) | (0.03) | (0.02) | (0.02) | (0.02) | (0.03) |
| Consolidation | −0.002 | −0.005 | −0.016 | −0.005 | −0.03 | −0.01 | −0.016 | −0.01 |
| (t-ratio) | (0.006) | (0.002) | (0.005) | (0.002) | (0.01) | (0.003) | (0.005) | (0.005) |
| Z * Consolidation | −0.004 | −0.008 | −0.006 | −0.009 | 0.02 | −0.003 | −0.001 | −0.02 |
| (t-ratio) | (0.007) | (0.012) | (0.004) | (0.007) | (0.02) | (0.006) | (0.001) | (0.01) |
| $\phi > 0$ if: | Never | Never | Never | Never | Never | Never | Never | Never |
| Z Variable | Interest rate Change | Exchange rate Change | Maastricht Dummy | Spread | Interest rate Change | Exchange rate Change | Maastricht Dummy | Spread |
| Estimation Method | Panel IV | Panel IV | Panel IV | Panel IV | Panel IV | Panel IV | Panel IV | Panel IV |
| | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) |
| $\text{Growth}_{t-1}$ | 0.86 | 0.86 | 0.84 | 0.85 | 0.86 | 0.85 | 0.83 | 0.84 |
| (t-ratio) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) |
| Consolidation | −0.02 | −0.006 | −0.006 | −0.01 | −0.02 | −0.02 | −0.02 | −0.02 |
| (t-ratio) | (0.006) | (0.002) | (0.002) | (0.003) | (0.006) | (0.006) | (0.005) | (0.007) |
| Z * Consolidation | −0.006 | 0.02 | 0.01 | 0.001 | −0.06 | 0.04 | 0.03 | 0.001 |
| (t-ratio) | (0.002) | (0.02) | (0.006) | (0.001) | (0.02) | (0.03) | (0.02) | (0.001) |
| $\phi > 0$ if: | $Z < -0.28$ | Never | Never | Never | $Z < -0.28$ | Never | Never | Never |

*Notes:* This Table presents the results of estimating equation (6) ($g_{it} = \alpha g_{it-1} + \beta D_{it} + \gamma Z_{it} D_{it} + \eta_i + \delta_t + v_{it}$) under the endogeneity assumption and for two alternative consolidation dummies (AA2010 and IMF). In this specification we regress the GDP growth rate on the consolidation dummy adding an interaction term between the dummy and eight alternative Z variables (Z * Consolidation). This interaction term allows the possibility of a positive effect of the fiscal consolidation episode for certain values of Z even if $\beta$ is negative. In particular, the overall effect of the fiscal adjustment is given by $\phi = \beta + \gamma Z_{it}$. The significance of the interaction term allows us to test if non-keynesian effects might appear under certain circumstances surrounding the consolidation episode as proposed in the literature (e.g. consumption channel if the debt of Government debt is high enough, investment channel if the adjustment is based on reducing public wages, lax monetary policy accompanying the consolidation...) The Anderson and Hsiao (1982) panel IV estimator is employed in all columns. Dependent variable is always the GDP growth rate. Standard errors reported in parentheses are clustered at the country level.

Table 3: Effect on Consumption, Investment and Unemployment

| | Investment | | Consumption | | Unemployment | |
|---|---|---|---|---|---|---|
| | AA2010 | IMF | AA2010 | IMF | AA2010 | IMF |
| Estimation Method | Panel IV | Panel IV | Panel IV | Panel IV | Panel IV | Panel IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\text{Growth}_{t-1}$ | 0.455 | 0.448 | 0.887 | 0.897 | 0.282 | 0.122 |
| | (0.075) | (0.070) | (0.018) | (0.024) | (0.053) | (0.107) |
| Consolidation | −0.055 | −0.105 | −0.002 | −0.021 | 0.039 | −0.006 |
| | (0.015) | (0.031) | (0.002) | (0.008) | (0.017) | (0.042) |
| Fixed Effects | YES | YES | YES | YES | YES | YES |
| AR(2) test p-value | 0.58 | 0.74 | 0.84 | 0.54 | 0.19 | 0.16 |
| Observations | 440 | 337 | 518 | 390 | 517 | 390 |
| Countries | 18 | 13 | 18 | 15 | 21 | 15 |

*Notes:* This Table presents the results of estimating equation (8), ($\text{Growth}^h_{it} = \alpha\ \text{Growth}^h_{it-1} + \beta\ \text{Consolidation}_{it} + \eta_i + \delta_t + v_{it}$ where $h$ is consumption, investment or unemployment) under the endogeneity assumption and for two alternative consolidation dummies (AA2010 and IMF). The Anderson and Hsiao (1982) panel IV estimator is employed in all the columns. AR(2) test p-value refers to the p-values from the Arellano and Bond's (1991) test for the lack of second order autocorrelation in the first differenced errors. Under this null hypothesis the instruments considered to address the endogeneity are valid because the errors in levels are not correlated as implied by assumption (3). Standard errors reported in parentheses are clustered at the country level.

Table 4: Characteristics of Sucessful Fiscal Consolidations

| | AA98 | A04 | G07 |
|---|---|---|---|
| Dependent variable is the successful consolidation dummy | | | |
| GDP growth | | 12.47 | |
| (t-ratio) | | (4.67) | |
| Government debt in $t-1$ | | $-0.02$ | $-0.23$ |
| (t-ratio) | | $(-0.17)$ | $(-1.58)$ |
| Deficit level in $t-1$ | | 0.34 | |
| (t-ratio) | | (0.31) | |
| Consolidation Size | 1.91 | 2.52 | 2.44 |
| (t-ratio) | (0.61) | (0.90) | (0.76) |
| $\Delta$ wage expenditures | $-0.03$ | $-0.41$ | 0.01 |
| (t-ratio) | $(-0.18)$ | $(-2.07)$ | (0.05) |
| $\Delta$ interest rate | | | $-0.01$ |
| (t-ratio) | | | $(-0.22)$ |
| $\Delta$ exchange rate | | | 0.01 |
| (t-ratio) | | | (0.32) |
| Output gap | | | 0.02 |
| (t-ratio) | | | (0.82) |
| $R^2$ | 0.01 | 0.29 | 0.06 |
| Obs. | 73 | 73 | 73 |

*Notes:* This Table presents the results from estimating three OLS regressions of the successful (in terms of debt reduction) consolidation dummy on the determinants suggested in Alesina and Ardagna (1998) [AA98], Ardagna (2004) [A04], and Giudice et al. (2007) [G07]. The change in wage expenditures is relative to the change in primary budget deficit (i.e. $\Delta$wage expenditures is one of the items in $\Delta item_{it}/\Delta Def_{it}$; an increase in this variable means that a larger share of the reduction of the primary deficit is obtained by cutting the public wage bill.).

Table 5: Characteristics of Successful Consolidations *via* Model Averaging

| | PIP | P. Mean | P. Std. |
|---|---|---|---|
| Output gap | 0.60 | 0.05 | 0.02 |
| ΔS.s. contributions | 0.42 | −0.58 | 0.30 |
| ΔOther taxes | 0.28 | −0.57 | 0.37 |
| ΔTransfers | 0.25 | 0.28 | 0.21 |
| GDP growth | 0.21 | 4.50 | 3.64 |
| ΔBusiness taxes | 0.20 | 0.31 | 0.27 |
| ΔGovernment investment | 0.16 | 0.15 | 0.16 |
| Deficit level | 0.15 | −1.36 | 1.61 |
| ΔNon-wage expenditures | 0.14 | −0.32 | 0.43 |
| Government debt level | 0.13 | −0.15 | 0.18 |
| ΔWage expenditures | 0.13 | −0.19 | 0.31 |
| ΔIndirect taxes | 0.12 | −0.13 | 0.26 |
| ΔInterest rate | 0.11 | −0.02 | 0.03 |
| ΔIncome taxes | 0.11 | −0.03 | 0.22 |
| Consolidation size | 0.11 | −1.87 | 3.84 |
| ΔSubsidies | 0.11 | −0.24 | 0.56 |
| ΔExchange rate | 0.10 | −0.01 | 0.02 |
| Prior Inclusion Probability | | 0.5 | |
| Number of models estimated | | 131,072 | |

*Notes:* PIP refers to the posterior inclusion probability of a particular regressors. Given the prior inclusion probability is equal for all the variables (i.e. 0.5), those variables with PIP higher than 0.5 are labeled as robust determinants of successful consolidations. All the regressors belonging to the public bill (e.g. ΔSubsidies, ΔIndirect taxes,...) are divided by the total change in the primary deficit to focus on the proportion of the adjustment which was due to a particular item as proxies of the composition. P. Mean refers to the posterior mean conditional on inclusion of a given regressor in the empirical model, which is a weighted average of model-specific coefficient estimates with weights given by the model-specific R-squares. P. Std. is the square root of the posterior variance which is a weighted average of model-specific variances also including the variance of the estimates across different models. The sample is formed by 73 country-year pairs in which a consolidation took place. The 131,072 estimated models come from all the possible combinations of the 17 regressors ($2^{17} = 131,072$).

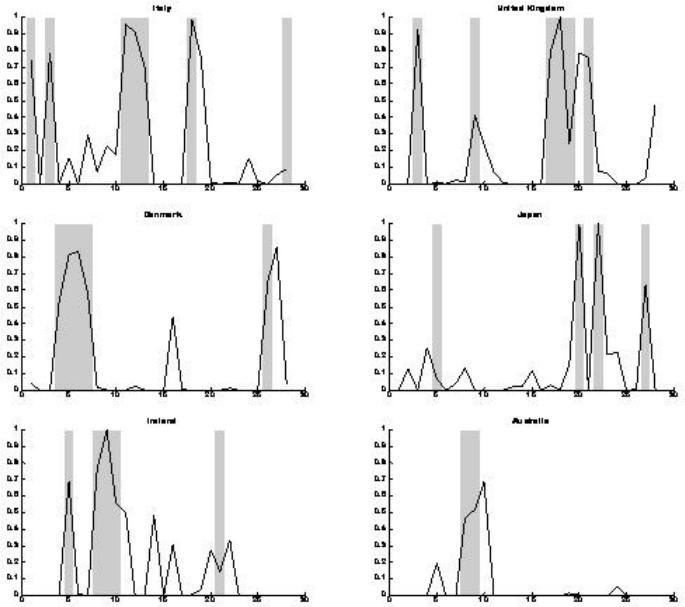
Table 6: Characteristics of Successful Consolidations: Alternative Definitions

| | Persistence Criterion | | | Expansionary Criterion | | |
|---|---|---|---|---|---|---|
| | PIP | P. Mean | P. Std. | PIP | P. Mean | P. Std. |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ΔWage expenditures | 0.72 | 0.38 | 0.15 | 0.10 | 0.05 | 0.22 |
| ΔIncome taxes | 0.34 | 0.19 | 0.11 | 0.11 | −0.07 | 0.16 |
| GDP growth | 0.17 | 2.13 | 2.00 | 0.10 | −0.51 | 2.81 |
| ΔIndirect taxes | 0.14 | −0.12 | 0.15 | 0.10 | −0.06 | 0.20 |
| ΔGovernment investment | 0.13 | −0.07 | 0.10 | 0.13 | −0.09 | 0.13 |
| ΔS.s. contributions | 0.13 | 0.13 | 0.19 | 0.16 | 0.28 | 0.27 |
| Output gap | 0.12 | 0.01 | 0.01 | 0.17 | −0.02 | 0.02 |
| ΔOther taxes | 0.12 | −0.12 | 0.25 | 0.10 | −0.10 | 0.31 |
| ΔTransfers | 0.12 | −0.06 | 0.12 | 0.14 | 0.14 | 0.16 |
| ΔBusiness taxes | 0.12 | −0.08 | 0.16 | 0.11 | −0.08 | 0.22 |
| ΔNon-wage expenditures | 0.11 | 0.12 | 0.24 | 0.12 | 0.25 | 0.34 |
| Government debt level | 0.10 | 0.04 | 0.11 | 0.10 | 0.04 | 0.15 |
| Deficit level | 0.10 | −0.29 | 0.82 | 0.19 | 1.46 | 1.20 |
| ΔInterest rate | 0.10 | 0.00 | 0.02 | 0.10 | 0.00 | 0.03 |
| ΔSubsidies | 0.10 | 0.06 | 0.31 | 0.10 | 0.11 | 0.46 |
| ΔExchange rate | 0.10 | 0.00 | 0.01 | 0.10 | 0.00 | 0.01 |
| Consolidation size | 0.10 | 0.11 | 2.19 | 0.26 | 4.84 | 3.20 |
| Prior Inclusion Probability | | 0.5 | | | 0.5 | |
| Number of models estimated | | 131,072 | | | 131,072 | |

*Notes:* PIP refers to the posterior inclusion probability of a particular regressors. Given the prior inclusion probability is equal for all the variables (i.e. 0.5), those variables with PIP higher than 0.5 are labeled as robust determinants of successful consolidations. All the regressors belonging to the public bill (e.g. ΔSubsidies, ΔIndirect taxes,...) are divided by the total change in the primary deficit to focus on the proportion of the adjustment which was due to a particular item as proxies of the composition. P. Mean refers to the posterior mean conditional on inclusion of a given regressor in the empirical model, which is a weighted average of model-specific coefficient estimates with weights given by the model-specific R-squares. P. Std. is the square root of the posterior variance which is a weighted average of model-specific variances also including the variance of the estimates across different models. The sample is formed by 73 country-year pairs in which a consolidation took place. The 131,072 estimated models come from all the possible combinations of the 17 regressors ($2^{17} = 131,072$).

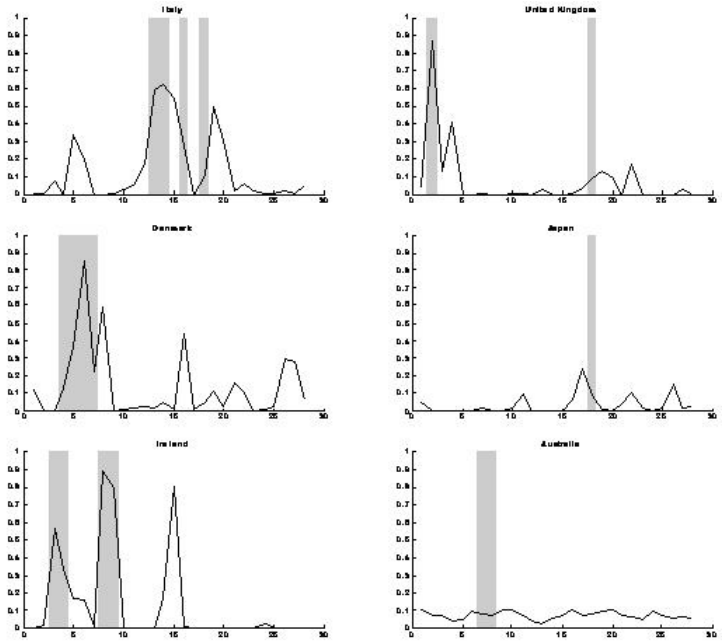
Figure 1: Fiscal Consolidations and Predicted Probabilities

AA2010 Dummy




This graph presents the consolidation episodes over the period 1980-2007 defined as in AA2010 (grey areas) and the precited probabilities from past economic information (solid line). Predicted probabilities are estimated from a country-by-country probit of the consolidation dummy on two lags of GDP growth and primary deficit.

IMF Dummy




This graph presents the consolidation episodes over the period 1980-2007 defined according to the IMF narrative approach (grey areas) and the precited probabilities from past economic information (solid line). Predicted probabilities are estimated from a country-by-country probit of the consolidation dummy on two lags of GDP growth and primary deficit.